

Two papers on
EFFICIENCY VERSUS PROTECTION
IN
RANDOMIZED RESPONSE DESIGNS
by
Harald Anderson



UNIVERSITY OF LUND AND
LUND INSTITUTE OF TECHNOLOGY
DEPARTMENT OF MATHEMATICAL STATISTICS

LUND, SWEDEN

1975

REPORT ON THE
PROGRESS OF THE
RESEARCH IN
THE
MATHEMATICAL SCIENCES
IN
THE
UNIVERSITY OF LEIDEN



This collection contains the following two papers

[A] EFFICIENCY VERSUS PROTECTION IN RANDOMIZED RESPONSE DESIGNS FOR
ESTIMATING PROPORTIONS

[B] EFFICIENCY VERSUS PROTECTION IN A GENERAL RANDOMIZED RESPONSE
MODEL

Department of Mathematical Statistics

Box 723

S-220 07 1300, Sweden

Paper [A]

EFFICIENCY VERSUS PROTECTION IN RANDOMIZED RESPONSE DESIGNS FOR
ESTIMATING PROPORTIONS

EFFICIENCY VERSUS PROTECTION
IN RANDOMIZED RESPONSE DESIGNS FOR
ESTIMATING PROPORTIONS

by
Herald Anderson

Ms. 1475/2

July 1975

Department of Mathematical Statistics

Box 725

S-220 07 LUND, Sweden

EFFICIENCY VERSUS PROTECTION

IN RANDOMIZED RESPONSE DESIGNS FOR

ESTIMATING PROPORTIONS

by

Harald Anderson

No. 1975:9

July 1975

EFFICIENCY VERSUS PROTECTION IN RANDOMIZED RESPONSE DESIGNS FOR ESTIMATING PROPORTIONS

by

Harald Anderson

1. INTRODUCTION AND SUMMARY

In sampling surveys that deal with embarrassing subjects refusal to answer and incorrect replies give rise to difficult problems. For example, it is often impossible to get satisfactory estimates of quantities such as the number of tax evaders, drug addicts and illegal abortions in a population.

An interviewing technique, which protects the privacy of the interviewees but still makes it possible to obtain unbiased estimates, is randomized response (RR for short). This technique, which was introduced by Warner (1965), is based on a very simple idea. Suppose, that in a population, one wants to estimate the proportion π having some sensitive characteristic A. The sampled persons are asked to draw and answer a question from a pack of cards with two types of questions,

Are you an A-individual?

and

Are you an A*-individual?,

A* denoting "not A". The proportions of cards are p and $1-p$, respectively, where p is a number $\neq 1/2$ determined by the statistician. Since the interviewer does not know which question is answered he cannot infer the attribute of the respondent. The proportion π can still be estimated, since the probability of a yes-answer is a linear function of π .

In Warner's RR-scheme both questions relate to the sensitive attribute A, and for psychological reasons this may be inconvenient: the respondents might suspect that they are being deceived. An RR-model, which is more flexible and also overcomes the above-mentioned drawback, is the method of "unrelated question" proposed by Simmons (Greenberg et al. (1969)). The respondent ran-

domly and with known probabilities p and $1-p$ chooses one of the questions

Are you an A-individual?

and

Are you a B-individual?,

where the attribute B is innocent, impossible to check and independent of A.

If the proportion of B-individuals, π_B , is known, π can be estimated from a single sample. If π_B is unknown, two samples with different mixtures of the above questions must be taken (see also Moors (1971), Lanke (1975a)).

The multi-proportion case is discussed by Abul-Ela et al. (1967) in essentially the same way as the single-proportion case is treated by Warner. To estimate the proportions of the classes $A_1, \dots, A_k$, they propose that $k-1$ samples with different mixtures of the questions

Are you an A_i -individual?, $i = 1, 2, \dots, k$,

are taken. The multi-proportion case has also been treated by Bourke and Dalenius (1973), who introduced a model requiring only a single sample.

Greenberg et al. (1971) and Eriksson (1973) have extended the RR-technique to the estimation of the mean of a population. In order to determine the frequency of illegal abortions in a population, the former asked the respondents to answer one of the randomly posed questions,

How many abortions have you had during your life-time?

and

If a woman has to work full-time to make a living, how many children do you think she should have?.

Eriksson proposed a so-called "given-answer model". The respondent then randomly answers according to either his true value or some known probability distribution. It is shown how this model can be used in different sampling schemes, and also that it is more flexible than earlier RR-models.

Note that many population parameters can be expressed as linear combinations of proportions; e.g. in the above example the average number of abortions per woman is $\mu = \sum i\pi_i$, where π_i is the proportion of women having had i abortions. The multi-proportion case thus covers more situations than might be anticipated.

Warner (1971) showed how different RR-models can be put into the framework of the general linear model. He also pointed out the possibility to contaminate official registers with fictive data before delivering them for more detailed analysis. See also Boruch (1972).

A common feature of all RR-models is that, somewhat contradictorily, one wants to know little about the single individuals of a sample but much about the population. It is natural to summarize the information about the population in terms of estimates of one or more population parameters and to measure their goodness by means of their variances. It is more difficult to formalize and quantify the protection of the respondents' privacy.

In the present paper the estimation of proportions and linear combinations of proportions is treated. First, in Section 2, a general RR-model is proposed. Individuals belonging to the classes A_i , $i = 1, \dots, k$, answer according to different but known probability distributions, and the proportions π_i , $i = 1, \dots, k$, are mixture parameters between these distributions. In Section 3 privacy aspects are discussed, and a measure of the protection based on conditional probabilities is introduced.

The single-proportion case is studied in Sections 4 and 5. By means of Cramér-Rao's inequality it is shown how the protection of privacy and the variance of the π -estimate are related. Some planning aspects are also given, and it is demonstrated that plans with two possible answers have certain optimal properties.

In Section 6 the results are generalized to the multi-proportion case. A more detailed consideration of the protection is then required. Moreover, the design problems are more delicate, since there is no natural single optimality criterion when several parameters are involved. Finally, in Section 7, we assume that the proportions of the different classes are functions of some unknown parameter θ , a situation occurring when a parameterized distribution is divided into classes.

2. THE MODEL

So far as we know, different RR-schemes have been analyzed in terms of the special randomizing devices being used. In this section a general RR-model for estimating proportions will be introduced permitting a unified approach to RR.

To begin with let us consider the single-proportion case: the population consists of A- and A*-individuals in the proportions π and $1-\pi$. In an RR-interview the answer Y is determined by the characteristic of the respondent, A or A*, and the outcome of a random experiment. This experiment is chosen by the statistician and he knows its random properties. Hence we have the general model:

A-individuals answer according to the known probability density $f_0(y)$ and

A*-individuals answer according to the known probability density $f_1(y)$. The densities $f_0(y)$ and $f_1(y)$ are called response densities and they hold with respect to some measure $d\mu(y)$; thus we include discrete as well as continuous response distributions. For instance, in Simmons's scheme we have with notations from Section 1,

$$f_0(\text{yes}) = P(\text{yes-answer from an A-individual}) = p + (1-p)\pi_B$$

and

$$f_1(\text{yes}) = P(\text{yes-answer from an A*-individual}) = (1-p)\pi_B.$$

If the RR-scheme allows repetitive interviews, the random variable Y is multi-dimensional.

It is also possible to realize any pair of response densities $f_0(y)$, $f_1(y)$ by means of e.g. a spinner with two scales, one for A-and one for A*-individuals, or a pack of cards with two answers on each card.

The answer from a randomly taken individual has the density

$$f_Y(y; \pi) = \pi f_0(y) + (1-\pi) f_1(y).$$

Hence π is a mixture parameter between two known densities.

In the multi-proportion case the population is composed by A_i -individuals in the proportion π_i , $i = 1, \dots, k$, $\sum \pi_i = 1$. The general RR-model is then,

A_i -individuals answer according to the known probability

density $f_i(y)$, $i = 1, \dots, k$,

and the answer from a randomly taken individual has the density

$$f_Y(y; \pi) = \sum_{i=1}^k \pi_i f_i(y),$$

where $\pi = (\pi_1, \dots, \pi_{k-1})^T$.

3. PRIVACY ASPECTS

3.1. General considerations

A main characteristic of an RR-interview is that the true value of the variable under study may not be inferred from an interview answer, at least if the true value is stigmatizing. The idea is that answer bias and answer refusal will decrease, because the protection of the privacy is thought to make the respondents more co-operative. In a planning situation the statistician has to judge whether the expected decrease of answer bias will balance the loss of information as well as the extra costs caused by RR.

A crucial point in the realization of an RR-interview is obviously to convince the respondent that a stigmatizing value cannot be revealed through the RR-interview. Here the behaviour of the interviewer, the explicit technical design of the randomizing device, and the statistical properties of that device play important roles. The first two factors are essentially non-statistical, and different techniques can be evaluated only by making comparative field studies. In the following we shall consider only the statistical properties of different RR-schemes. These properties determine the protection provided by the RR-scheme and hence the risk to participate in an interview for a rational respondent. Moreover, when using stochastic transformations to contaminate a register, the probabilistic aspects of the randomizing device are decisive.

Earlier authors have considered how the willingness to co-operate in an RR-interview depends on the design-parameters used in their special variants of RR. For example, Greenberg et al. (1969) and Moors (1971) discuss how the proportion p of the sensitive question and the probability π_B of a yes-answer to the unrelated question should be taken in Simmons's model. This approach is profitable in a practical situation if the type of RR-scheme already is determined, but it makes it difficult to evaluate and compare the statistical properties of different typed of schemes. More general aspects have been given by Bourke (1974). He advocates symmetric RR-schemes,

meaning that any possible answer might come from both an A- and an A*-individual. Hence there is no answer that completely clears the respondent from the suspicion of being an A-individual (or an A*-individual), and this fact should influence the interviewee to answer correctly.

Lanke (1975a, 1975b) has studied the qualities of some RR-schemes with two possible answers, yes and no, for estimating a single proportion π . He proposes that the conditional probabilities of being an A-individual given the answers yes and no, $P(A|yes;\pi)$ and $P(A|no;\pi)$, should be used to measure the protection associated with these answers. Lanke assumes that only the characteristic A is embarrassing, and he uses the quantity $\max(P(A|yes;\pi), P(A|no;\pi))$ as a measure of the protection provided by an RR-scheme.

In the following sections we shall study the privacy aspects in terms of the general model of Section 2. In the next subsection we shall for this purpose introduce a general measure of the protection.

3.2. A statistical measure of the protection

The idea of using conditional probabilities to measure the protection can be generalized to RR-schemes with several possible answers. Let, as in Section 2, an RR-answer be represented by the random variable Y and let Ω_y be the set of possible RR-answers. This set is determined by the specific RR-scheme considered.

First, we discuss the single-proportion case; the multi-proportion case will be treated in Section 6. Given the answer $Y = y$ the probability that a randomly taken respondent has the characteristic A is $P(A|y;\pi)$, and the probability that he has the characteristic A* is $P(A^*|y;\pi) = 1 - P(A|y;\pi)$. These probabilities measure the protection associated with the answer $Y = y$, and in the sequel they will often be called risks of suspicion. From Bayes' theorem we immediately have

$$P(A|y;\pi) = \frac{\pi f_0(y)}{\pi f_0(y) + (1-\pi) f_1(y)}, \quad y \in \Omega_y,$$

and

$$P(A^*|y;\pi) = \frac{(1-\pi)f_1(y)}{\pi f_0(y) + (1-\pi)f_1(y)}, \quad y \in \Omega_y,$$

where $f_0(y)$ and $f_1(y)$ are the response densities of the RR-scheme.

It is clear that a non-informative answer y (or no answer at all) keeps the risks of suspicion unchanged:

$$P(A|y;\pi) = 1 - P(A^*|y;\pi) = \pi.$$

On the other hand, if the interviews are direct, the risks of suspicion will either be one or zero. A non-trivial RR-answer y will involve risks of suspicion between these limits. Either the answer will increase the probability that the respondent is an A-individual, $P(A|y;\pi) > \pi$, or it will increase the probability that he is an A*-individual, $P(A^*|y;\pi) > 1-\pi$, compared to the a priori situation. The larger the absolute difference between $P(A|y;\pi)$ and π is, the more informative is the answer y . In fact, we shall in the following sections see that the quantity $|P(A|y;\pi) - \pi|$ in a decisive way affects the efficiency of an RR-scheme to estimate π .

So far we have considered the protection associated with a given answer y . To get an overall measure of the protection provided by an RR-scheme, all the possible answers must be accounted for. In the next subsection we shall for this purpose consider the quantities $\max_y P(A|y;\pi)$ and $\max_y P(A^*|y;\pi)$.

3.3. Planning aspects

An essential task in designing an RR-survey is to keep the risks of suspicion down, because they influence the willingness of the interviewee to co-operate. It is natural to require that the risks of suspicion $P(A|y;\pi)$, $y \in \Omega_y$, be restricted by an upper bound $k_2 < 1$, if the characteristic A is regarded as embarrassing, and, similarly, that the risks of suspicion $P(A^*|y;\pi)$, $y \in \Omega_y$, be restricted by an upper bound $1-k_1 < 1$, if A* is embarrassing. This leads us to the conditions

$$\max_y P(A|y;\pi) \leq k_2$$

and

$$\max_y P(A^*|y;\pi) \leq 1-k_1$$

or, equivalently,

$$(3.1) \quad k_1 \leq P(A|y;\pi) \leq k_2, \quad y \in \Omega_y.$$

Condition (3.1) guarantees the interviewee a minimal protection against being revealed whatever his answer will be, thus promoting a high participation. Another important argument for requiring uniform protection over the set of possible answers is the fact that the respondent is free to refuse to answer or to lie when he knows the outcome of the random experiment determining his answer; if a possible answer y involves too high a risk of suspicion, the respondent will probably give an evasive answer and thus contribute to a biased estimate.

Very often, when planning a statistical survey, one should preferably know the parameter to be estimated in order to construct a good plan. Of course, this is impossible, and so the design must be based on a guess about the population. In this case the risks of suspicion depend on the unknown proportion π and Condition (3.1) cannot be applied directly. The traditional way of solving this problem is to fix a π , believed to be approximately correct, and determine an optimal RR-plan subject to (3.1) for this particular π .

The fixed π reflects the statistician's a priori knowledge about the population. However, in RR-interviews the respondents judge the risks of suspicion and hence their a priori probabilities of taking an A-individual should be considered. These probabilities are unknown and vary from individual to individual. Moreover, after the survey has been carried out, the estimate of π and the risks of suspicion might be different from what was a priori expected. A possible solution of these problems is to choose a design such that Condition (3.1) is fulfilled for all π in some interval (π_1, π_2) believed to contain the true value of π :

$$(3.2) \quad k_1 \leq P(A|y;\pi) \leq k_2, \quad y \in \Omega_y, \quad \pi_1 \leq \pi \leq \pi_2.$$

This means that the interviewees are guaranteed that the risks of suspicion are

never larger than k_2 and $1-k_1$, respectively, regardless of the answer y and regardless of which value π assumes in the interval (π_1, π_2) . From a Bayesian point of view the risk of suspicion $P(A|y; \pi)$ will assume values in the interval (k_1, k_2) for every $y \in \Omega_y$ and every a priori probability π in the interval (π_1, π_2) .

In the next two sections both Condition (3.1) and Condition (3.2) will be applied to the single-proportion case and in Section 6, the conditions are generalized to the multi-proportion case.

4. TWO-ANSWER MODEL FOR ESTIMATING A SINGLE PROPORTION

4.1. Model and variance of π -estimate

Let the population consist of A- and A*-individuals in the proportions π and $1-\pi$, respectively. Consider a one-sample RR-scheme with two possible answers, yes and no (e.g. Warner's or Simmons's model). A given randomizing mechanism determines the response probabilities from A- and A*-individuals,

$$P(\text{yes}|A) = 1 - P(\text{no}|A)$$

and

$$P(\text{yes}|A^*) = 1 - P(\text{no}|A^*).$$

Suppose we take a simple random sample of n individuals with replacement, n_1 of whom answer yes. The probability of a yes-answer is

$$(4.1) \quad P(\text{yes}; \pi) = \pi P(\text{yes}|A) + (1-\pi)P(\text{yes}|A^*).$$

Hence the maximum likelihood estimate $\hat{\pi}$ of π is obtained from the relation

$$n_1/n = \hat{\pi}P(\text{yes}|A) + (1-\hat{\pi})P(\text{yes}|A^*);$$

thus

$$\hat{\pi} = \frac{n_1/n - P(\text{yes}|A^*)}{P(\text{yes}|A) - P(\text{yes}|A^*)}.$$

The ML-estimate $\hat{\pi}$ is an efficient estimate (compare the end of Section 5.1) with

$$E(\hat{\pi}) = \pi$$

and

$$(4.2) \quad V(\hat{\pi}) = \frac{P(\text{yes}; \pi)(1-P(\text{yes}; \pi))}{n(P(\text{yes}|A) - P(\text{yes}|A^*))^2}.$$

The suspicion that a respondent has the attribute A is after a yes- and a no-answer, respectively,

$$(4.3) \quad \begin{aligned} P(A|\text{yes}; \pi) &= \frac{\pi P(\text{yes}|A)}{P(\text{yes}; \pi)} = \frac{\pi P(\text{yes}|A)}{\pi P(\text{yes}|A) + (1-\pi)P(\text{yes}|A^*)}, \\ P(A|\text{no}; \pi) &= \frac{\pi P(\text{no}|A)}{P(\text{no}; \pi)} = \frac{\pi P(\text{no}|A)}{\pi P(\text{no}|A) + (1-\pi)P(\text{no}|A^*)}. \end{aligned}$$

The variance $V(\hat{\pi})$ can now be expressed in terms of the risks of suspicion $P(A|\text{yes}; \pi)$ and $P(A^*|\text{no}; \pi)$. From the relations (4.3) we have

$$P(\text{yes}|A) = \frac{P(A|\text{yes};\pi)(\pi - P(A|\text{no};\pi))}{\pi(P(A|\text{yes};\pi) - P(A|\text{no};\pi))}$$

and

$$P(\text{yes}|A^*) = \frac{(\pi - P(A|\text{no};\pi))(1 - P(A|\text{yes};\pi))}{(1 - \pi)(P(A|\text{yes};\pi) - P(A|\text{no};\pi))}.$$

Inserting these expressions in (4.1) we get

$$P(\text{yes};\pi) = \frac{\pi - P(A|\text{no};\pi)}{P(A|\text{yes};\pi) - P(A|\text{no};\pi)}.$$

Hence the variance $V(\hat{\pi})$ can be written

$$nV(\hat{\pi}) = \frac{\pi^2(1-\pi)^2}{(\pi - P(A|\text{no};\pi))(P(A|\text{yes};\pi) - \pi)}$$

or, after setting $P(A|\text{no};\pi) = 1 - P(A^*|\text{no};\pi)$,

$$nV(\hat{\pi}) = \pi(1-\pi) \cdot \frac{1}{\left[\frac{P(A|\text{yes};\pi)}{\pi} - 1\right]\left[\frac{P(A^*|\text{no};\pi)}{1-\pi} - 1\right]}.$$

Note that we are free to choose the response probabilities $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$. In the sequel we shall assume that $P(\text{yes}|A) > P(\text{yes}|A^*)$ making a yes-answer increase and a no-answer decrease the probability of being an A-individual:

$$P(A|\text{no};\pi) < \pi < P(A|\text{yes};\pi).$$

To summarize we have the following

THEOREM 1 The minimum variance estimate $\hat{\pi}$ of a two-answer model has the variance

$$V(\hat{\pi}) = \Psi(\pi) \cdot \pi(1-\pi)/n,$$

where

$$\Psi(\pi) = \frac{1}{\left[\frac{P(A|\text{yes};\pi)}{\pi} - 1\right]\left[\frac{P(A^*|\text{no};\pi)}{1-\pi} - 1\right]}.$$

From Theorem 1 it follows that, compared with an ordinary binomial situation, the variance increases with the factor $\Psi(\pi) \geq 1$. Note also that, for given π , all two-answer plans with the same risks of suspicion $P(A|\text{yes};\pi)$ and $P(A^*|\text{no};\pi)$ are equally effective to estimate π irrespectively of the randomizing procedure. Different two-answer RR-schemes differ from a statistical point of view only in that they permit different

response probabilities $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$, hence generating different risks of suspicion $P(A|\text{yes};\pi)$ and $P(A^*|\text{no};\pi)$.

EXAMPLE 1 Simmons's model. The interviewee answers with probability p the question

Are you an A-individual?

and with probability $1-p$ the question

Are you a B-individual?.

Assuming that the characteristic B is independent of A and that $P(B) = \pi_B$ is known, we have $P(\text{yes}|A) = p + \pi_B(1-p)$ and $P(\text{yes}|A^*) = \pi_B(1-p)$. Note that any values of $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$ with $P(\text{yes}|A) > P(\text{yes}|A^*)$ can be generated by an appropriate choice of p and π_B . We get

$$P(A|\text{yes};\pi) = \frac{\pi(p + \pi_B(1-p))}{\pi(p + \pi_B(1-p)) + (1-\pi)\pi_B(1-p)},$$

$$P(A^*|\text{no};\pi) = \frac{(1-\pi)(1 - \pi_B(1-p))}{\pi(1-p - \pi_B(1-p)) + (1-\pi)(1 - \pi_B(1-p))}$$

and from Theorem 1

$$nV(\hat{\pi}) = \pi(1-\pi) \left[\frac{p + \pi_B(1-p)}{\pi(p + \pi_B(1-p)) + (1-\pi)\pi_B(1-p)} - 1 \right]^{-1}$$

$$\cdot \left[\frac{1 - \pi_B(1-p)}{\pi(1-p - \pi_B(1-p)) + (1-\pi)(1 - \pi_B(1-p))} - 1 \right]^{-1} =$$

$$= \dots = (\pi p + \pi_B - \pi_B p)(1 - \pi p - \pi_B + \pi_B p) / p^2 = P(\text{yes};\pi)P(\text{no};\pi) / p^2.$$

This result of course also follows directly from (4.3).

4.2. Planning aspects

When choosing the two design parameters $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$ two conflicting aspects must be considered. The risks of suspicion $P(A|\text{yes})$ and $P(A^*|\text{no})$ (in the following we often omit the argument π when no misunderstanding can arise) should be kept down and the variance of the π -estimate should be small. According to the privacy discussion of Section 3 a reasonable protection is obtained by restricting the risks uniformly for all possible answers and all π in some interval (π_1, π_2) . This means for the two-answer

plan that

$$(4.4) \quad k_1 \leq P(A|\text{no};\pi) < P(A|\text{yes};\pi) \leq k_2, \quad \pi_1 \leq \pi \leq \pi_2,$$

where the limit risks k_1 and k_2 and the interval limits π_1 and π_2 are chosen in every special situation in such a way that

$$k_1 < \pi_1 < \pi_2 < k_2.$$

Using the relations (4.3) the condition (4.4) can be written

$$k_1 \leq \frac{\pi P(\text{no}|A)}{\pi P(\text{no}|A) + (1-\pi)P(\text{no}|A^*)} \leq \frac{\pi P(\text{yes}|A)}{\pi P(\text{yes}|A) + (1-\pi)P(\text{yes}|A^*)} \leq k_2,$$

$$\pi_1 \leq \pi \leq \pi_2,$$

from which

$$\frac{P(\text{yes}|A^*)}{P(\text{yes}|A)} \geq \frac{\pi(1-k_2)}{k_2(1-\pi)},$$

$$\pi_1 \leq \pi \leq \pi_2.$$

$$\frac{P(\text{no}|A^*)}{P(\text{no}|A)} \leq \frac{\pi(1-k_1)}{k_1(1-\pi)}.$$

Since the right members of the above inequalities are monotone functions of π we have the equivalent inequalities

$$(4.5) \quad \frac{P(\text{yes}|A^*)}{P(\text{yes}|A)} \geq \frac{\pi_2(1-k_2)}{k_2(1-\pi_2)},$$

$$\frac{P(\text{no}|A^*)}{P(\text{no}|A)} \leq \frac{\pi_1(1-k_1)}{k_1(1-\pi_1)}.$$

Returning to the variance $V(\hat{\pi})$, the design parameters $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$ should be taken subject to the conditions (4.5) so that $V(\hat{\pi})$ is minimized. But

$$V(\hat{\pi}) = \Psi(\pi) \cdot \pi(1-\pi)/n,$$

where

$$\begin{aligned} (\Psi(\pi))^{-1} &= \left[\frac{P(A|\text{yes})}{\pi} - 1 \right] \left[\frac{P(A|\text{no})}{1-\pi} - 1 \right] = \\ &= \left[\frac{P(\text{yes}|A)}{\pi P(\text{yes}|A) + (1-\pi)P(\text{yes}|A^*)} - 1 \right] \left[\frac{P(\text{no}|A)}{\pi P(\text{no}|A) + (1-\pi)P(\text{no}|A^*)} - 1 \right] \\ &= \left[\frac{1}{\pi + (1-\pi) \frac{P(\text{yes}|A^*)}{P(\text{yes}|A)}} - 1 \right] \left[\frac{1}{1-\pi + \pi \frac{P(\text{no}|A)}{P(\text{no}|A^*)}} - 1 \right]. \end{aligned}$$

We see that $\Psi(\pi)^{-1}$ is maximized and hence $V(\hat{\pi})$ is minimized independently of π when

$$\frac{P(\text{yes}|A^*)}{P(\text{yes}|A)} = \frac{\pi_2(1-k_2)}{k_2(1-\pi_2)}$$

and

$$\frac{P(\text{no}|A^*)}{P(\text{no}|A)} = \frac{\pi_1(1-k_1)}{k_1(1-\pi_1)} .$$

Solving for $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$ we have the optimal values

$$(4.6) \quad \begin{aligned} P(\text{yes}|A) &= \frac{(1-\pi_2)k_2(\pi_1-k_1)}{\pi_1(1-\pi_2)k_2(1-k_1) - \pi_2(1-\pi_1)k_1(1-k_2)} , \\ P(\text{yes}|A^*) &= \frac{\pi_2(1-k_2)(k_1-\pi_1)}{\pi_1(1-\pi_2)k_2(1-k_1) - \pi_2(1-\pi_1)k_1(1-k_2)} . \end{aligned}$$

Another approach to the design problem is to choose some appropriate values of $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$ and control the risks of suspicion, $P(A|\text{yes};\pi)$ and $P(A^*|\text{no};\pi)$, by illustrating them in a diagram as functions of π . Let, for instance, $P(\text{yes}|A) = P(\text{no}|A^*) = 0.80$. Then

$$P(A|\text{yes};\pi) = \frac{\pi}{\pi + 0.25(1-\pi)} ,$$

$$P(A^*|\text{no};\pi) = \frac{1-\pi}{0.25\pi + (1-\pi)} .$$

These probabilities are illustrated in Figure 1.

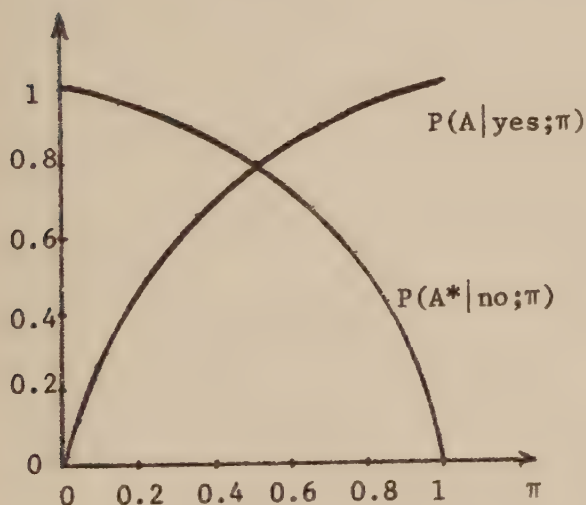


Figure 1. The risks of suspicion $P(A|\text{yes};\pi)$ and $P(A^*|\text{no};\pi)$ as functions of π for a plan with the response probabilities $P(\text{yes}|A) = P(\text{no}|A^*) = 0.80$.

An important special case occurs when the attribute A^* is not regarded as stigmatizing. This means that $P(A^*|\text{no})$ may take the value 1 implying that $k_1 = 0$ in (4.4) (compare e.g. Lanke (1975a)). Hence from (4.6) the

optimal response probabilities are

$$P(\text{yes}|A) = 1,$$

$$P(\text{yes}|A^*) = \pi_2(1-k_2)/(k_2(1-\pi_2)),$$

i.e. an A-individual should always answer yes. The A-individuals are protected by the possibility of yes-answers from A*-individuals.

Very often the characteristic A is rare, meaning that $P(A) = \pi$ is small. For small values of π we have

$$\frac{P(A|\text{yes};\pi)}{\pi} = \frac{P(\text{yes}|A)}{\pi P(\text{yes}|A) + (1-\pi)P(\text{yes}|A^*)} \approx \frac{P(\text{yes}|A)}{P(\text{yes}|A^*)},$$

which is independent of π . Since for small π

$$P(\text{yes})P(\text{no}) \approx P(\text{yes}|A^*)P(\text{no}|A^*)$$

we also have, according to (4.2),

$$nV(\hat{\pi}) \approx \frac{P(\text{yes}|A^*)P(\text{no}|A^*)}{(P(\text{yes}|A) - P(\text{yes}|A^*))^2},$$

which again is a constant independent of π . For example, when $P(\text{yes}|A) = P(\text{no}|A^*) = 0.80$, the approximations differ from the correct value with less than 10% for $\pi \leq 0.04$.

5. ARBITRARY RESPONSE DISTRIBUTIONS FOR ESTIMATING A SINGLE PROPORTION

5.1. Model and variance bound

As before we have a population composed by A- and A*-individuals in the proportions π and $1-\pi$ respectively. In Section 4 we studied RR-schemes with two possible answers. We shall now assume that A- and A*-individuals answer according to the arbitrary response densities $f_0(y)$ and $f_1(y)$. The observable answer from a randomly taken individual has the density

$$(5.1) \quad f_Y(y; \pi) = \pi f_0(y) + (1-\pi)f_1(y), \quad 0 < \pi < 1,$$

and the risks of suspicion are

$$(5.2) \quad \begin{aligned} P(A|y; \pi) &= \pi f_0(y) / f_Y(y; \pi), \\ P(A^*|y; \pi) &= (1-\pi) f_1(y) / f_Y(y; \pi), \end{aligned} \quad y \in \Omega_y.$$

Let $P(A|Y; \pi)$ be the random risk of suspicion, i.e. the random variable which takes the value $P(A|y; \pi)$ with probability $f_Y(y; \pi)$. Then we have

$$(5.3) \quad \begin{aligned} E(P(A|Y; \pi)) &= \pi, \\ V(P(A|Y; \pi)) &= E(P(A|Y; \pi) - \pi)^2. \end{aligned}$$

Hence $V(P(A|Y; \pi))$ is the expected square of the quantity $|P(A|Y; \pi) - \pi|$, mentioned in Section 3.2.

We shall now compute Fisher's information I about π from an observation of Y . Since $f_Y(y; \pi)$ is a regular density we have by definition

$$I = E\left[\frac{\partial}{\partial \pi} \log f_Y(Y; \pi)\right]^2$$

and by (5.1)

$$I = E\left[\frac{f_0(Y) - f_1(Y)}{f_Y(Y; \pi)}\right]^2$$

or by (5.2) and (5.3)

$$I = E\left[\frac{P(A|Y)}{\pi} - \frac{P(A^*|Y)}{1-\pi}\right]^2 = \frac{V(P(A|Y))}{\pi^2(1-\pi)^2}.$$

It now follows from the Cramér-Rao inequality that any unbiased estimate π^* of π has a variance not less than $(nI)^{-1}$. Summing up these results we have proved

THEOREM 2 For any unbiased estimate π^* of π based on n independent observations

$$nV(\pi^*) \geq 1/I,$$

where

$$I = E \left[\frac{f_0(Y) - f_1(Y)}{f_Y(Y; \pi)} \right]^2$$

or, alternatively,

$$I = \frac{V(P(A|Y; \pi))}{\pi^2 (1-\pi)^2}.$$

Since we are dealing with a regular case, the minimum variance bound is asymptotically attained by the maximum likelihood estimate. In sample surveys the sample size n is often large, and hence the variance may be determined approximately using this bound.

Boes (1966) has shown that the minimum variance bound is attained for a mixture parameter when n is finite if and only if the sample space can be partitioned into two domains Λ_1 and Λ_2 , such that

$$f_0(y)/f_1(y) \text{ is constant for } y \in \Lambda_i, i = 1, 2,$$

which is equivalent to

$$P(A|y) \text{ is constant for } y \in \Lambda_i, i = 1, 2.$$

This means that the minimum variance bound is attained for small samples only when $P(A|Y)$ assumes just two values, including as a special case the two-answer situation of Section 4.1.

4.2. Planning aspects

To design an RR-plan for estimating a proportion π means to choose two response densities $f_0(y)$ and $f_1(y)$ for A^- and A^* -individuals, respectively. Then two aspects must be considered; the risks of suspicion $P(A|y; \pi)$ and $P(A^*|y; \pi)$ should be kept low and the variance of the π -estimate should be small. In Section 3 we discussed the privacy aspects and found that a reasonable protection of the interviewee is obtained by restricting the risks uniformly for all answers y and for all π in some interval (π_1, π_2) :

$$(5.4) \quad k_1 \leq P(A|y;\pi) \leq k_2, \quad y \in \Omega_y, \quad \pi_1 \leq \pi \leq \pi_2.$$

We will now look for response densities, which, subject to the above restriction, yield a π -estimate with smallest possible variance. It will be shown that the solution to this problem is a plan with essentially two possible answers. To begin with, we assume that π is fixed. Then we have

THEOREM 3 Suppose the risk of suspicion $P(A|y;\pi)$ is to take values in the interval (k_1, k_2) for a fixed π and for every possible answer y . Then the most efficient plan to estimate π has essentially two possible answers.

PROOF From the inequality of Theorem 2 we see that large values of $V(P(A|Y))$ are preferable. Hence we want to find a plan such that $V(P(A|Y))$ is maximum subject to Condition (5.4) for a fixed π . This problem is equivalent to choosing a probability distribution of a random variable Z taking values in a given interval (k_1, k_2) with given mean π , such that its variance is as large as possible. It is immediately seen that the probability mass of Z should then be divided between the endpoints k_1 and k_2 in the proportions q and $1-q$ where $qk_1 + (1-q)k_2 = \pi$.

Returning to the original problem, we should use a design such that $P(A|Y)$ assumes the values k_1 and k_2 with probabilities q and $1-q$. For simplicity we let Y assume only two values, namely no and yes, corresponding to k_1 and k_2 , respectively. From the end of Section 4.1, we also have that the variance of the ML-estimate of a two-answer plan attains the minimum variance bound of Theorem 2. This completes the proof. $\square$

Now consider the case when condition (5.4) is to hold for all $\pi \in (\pi_1, \pi_2)$. From (5.1) and (5.2) we have

$$P(A|y;\pi) = \frac{\pi}{\pi + (1-\pi)f_1(y)/f_0(y)}.$$

Hence from (5.4)

$$\max_{\pi_1 \leq \pi \leq \pi_2} P(A|y;\pi) = \frac{\pi_2}{\pi_2 + (1-\pi_2)f_1(y)/f_0(y)} \leq k_2$$

and

$$\min_{\pi_1 \leq \pi \leq \pi_2} P(A|y;\pi) = \frac{\pi_1}{\pi_1 + (1-\pi_1)f_1(y)/f_0(y)} \geq k_1, \quad y \in \Omega_y,$$

or

$$(5.5) \quad \frac{\pi_2(1-k_2)}{(1-\pi_2)k_2} \leq \frac{f_1(y)}{f_0(y)} \leq \frac{\pi_1(1-k_1)}{(1-\pi_1)k_1}, \quad y \in \Omega_y.$$

Consider once again a fixed π . The response densities $f_0(y)$ and $f_1(y)$ are to be chosen subject to Condition (5.5). This condition is equivalent to

$$(5.6) \quad \frac{\pi}{\pi + (1-\pi)a_1} \leq P(A|y;\pi) \leq \frac{\pi}{\pi + (1-\pi)a_2}, \quad y \in \Omega_y,$$

where $a_i = \pi_i(1-k_i)/(k_i(1-\pi_i))$, $i = 1, 2$, are the limits in Condition (5.5). But from Theorem 3 we have that the optimal plan in this situation is a plan where $P(A|Y;\pi)$ assumes only the two boundary values in (5.6), that is when $f_1(y)/f_0(y)$ only assumes the two values a_1 and a_2 . This result is independent of π . Hence we have the following

THEOREM 4 Suppose the risk of suspicion $P(A|y;\pi)$ is to take values in the interval (k_1, k_2) for every possible answer and all $\pi \in (\pi_1, \pi_2)$. Then the most efficient plan to estimate π has essentially two possible answers.

REMARK 1 The explicit values of the design parameters $P(\text{yes}|A)$ and $P(\text{yes}|A^*)$ can be found in Section 4.2, where the two-answer case is studied in detail.

REMARK 2 The theorem says that every interviewee should be exposed to a boundary risk, thus contributing maximum information within the limits of protection.

6. ESTIMATION OF TWO OR MORE PROPORTIONS

6.1. Model and variance bound

The discussion will now be extended to the case of several unknown proportions. Let π_i be the probability that a randomly taken individual has the attribute A_i , $i = 1, 2, \dots, k$; $\sum \pi_i = 1$. The known response densities are $f_i(y)$, $i = 1, 2, \dots, k$, with respect to some measure $du(y)$, and the observable answer Y from a randomly taken individual has the density

$$f_Y(y; \pi) = \sum_1^k \pi_i f_i(y) = \sum_1^{k-1} \pi_i (f_i(y) - f_k(y)) + f_k(y),$$

where $\pi_i > 0$, $i = 1, 2, \dots, k$.

Fisher's information matrix I has the elements

$$I_{ij} = E \left[\frac{\partial \log f_Y(Y; \pi)}{\partial \pi_i} \cdot \frac{\partial \log f_Y(Y; \pi)}{\partial \pi_j} \right] = \text{Cov}(U_i, U_j), \quad i, j = 1, \dots, k-1,$$

where

$$U_i = \frac{f_i(Y) - f_k(Y)}{f_Y(Y; \pi)} = \frac{P(A_i | Y)}{\pi_i} - \frac{P(A_k | Y)}{\pi_k}, \quad i = 1, \dots, k-1.$$

Note that

$$\pi_i f_i(y) = P(A_i | y) f_Y(y; \pi), \quad i = 1, \dots, k.$$

The parameter $\pi = (\pi_1, \dots, \pi_{k-1})^T$ can be estimated only if the inverse I^{-1} exists. This occurs if and only if the random variable $U = (U_1, \dots, U_{k-1})$ does not with probability one lie in some subspace with $j < k-1$ dimensions. We have (compare Teicher (1963)):

THEOREM 5 The information matrix I of the $(k-1)$ -dimensional mixture parameter π is singular if and only if there exist constants c_i , $i = 1, \dots, k$, not all equal to zero, such that $\sum_1^k c_i f_i(y) = 0$ with probability one.

PROOF (a) Necessity.

If I is singular, there exist constants $c_i^!$, $i = 1, \dots, k-1$, not all equal to zero, such that $\sum_1^{k-1} c_i^! U_i = 0$. This implies that $\sum_1^{k-1} c_i^! (f_i(y) - f_k(y)) = 0$, which is of the form $\sum_1^k c_i f_i(y) = 0$.

(b) Sufficiency.

$\sum_1^k c_i f_i(y) = 0$ implies that $\sum_1^k c_i = 0$, since $f_i(y)$, $i = 1, \dots, k$, are

probability densities. Hence we have that $\sum_1^{k-1} c_i (f_i(y) - f_k(y)) = 0$ and $\sum_1^{k-1} c_i U_i = \sum_1^{k-1} c_i (f_i(y) - f_k(y)) / f_Y(y; \pi) = 0$. This holds with probability one and hence I is singular. $\square$

Consider now a linear combination $\lambda = \sum_1^{k-1} t_i \pi_i$ of the π_i 's. For an unbiased estimate λ^* based on a sample of n independent observations we have, assuming that I^{-1} exists,

$$(6.1) \quad nV(\lambda^*) \geq C$$

where

$$C = \sum t_i t_j (I^{-1})_{ij};$$

when $n \rightarrow \infty$ the Cramér-Rao bound C is asymptotically attained by the ML-estimate.

In Section 6.3 we shall need this bound in another form. This form also gives us a more comprehensible expression similar to that in the single-proportion case. Perform a linear transformation of U ,

$$X = TU,$$

where the $(k-1) \times (k-1)$ matrix $T = (t_{ij})$ has the elements $t_1, \dots, t_{k-1}$ in the first row; without loss of generality we may assume that $\sum t_i^2 = 1$. Moreover, we choose the other elements in T so that T becomes orthogonal. The moment matrix of X is given by

$$W = TIT^T$$

making

$$I^{-1} = T^T W^{-1} T,$$

where $W^{-1} = (w_{ij}^{-1})$ is the inverse of W . The Cramér-Rao bound for λ^* can then be written (see Cramér (1946) p. 305)

$$(6.2) \quad C = (t_1, \dots, t_{k-1}) T^T W^{-1} T \begin{bmatrix} t_1 \\ \vdots \\ t_{k-1} \end{bmatrix} = w_{11}^{-1} = \\ = \det W_{11} / \det W = 1 / \min_{\beta_i} V(X_1 - \sum_2^{k-1} \beta_i X_i),$$

where W_{11} is the matrix obtained by cancelling the first row and the first column of W . If we let ρ denote the multiple correlation coefficient

between $X_1 = \sum t_i U_i$ and $X_2, \dots, X_{k-1}$, the inequality (6.1) becomes

$$nV[(\sum t_i \pi_i)^*]V[\sum t_i \left[\frac{P(A_i|Y)}{\pi_i} - \frac{P(A_k|Y)}{\pi_k} \right]](1-\rho^2) \geq 1.$$

6.2. Planning aspects

The planning problems become more complex, when we deal with several proportions. Since the two-parameter case shows the essential characteristics we shall mainly consider that case. There are then three attributes, A_1, A_2, A_3 , and an RR-plan is determined by the three response densities $f_1(y), f_2(y), f_3(y)$. As in the single-proportion case any triple of response densities can be realized by, for example, a spinner with three scales. The protection provided is measured by the risks of suspicion $P(A_1|y;\pi)$, $P(A_2|y;\pi)$ and $P(A_3|y;\pi) = 1 - P(A_1|y;\pi) - P(A_2|y;\pi)$, $y \in \Omega_y$, and the efficiency to estimate $\pi = (\pi_1, \pi_2)^T$ will be measured by means of the corresponding information matrix I (see Section 6.1).

We wish to choose an RR-plan, i.e. a triple $f_1(y), f_2(y), f_3(y)$, such that the risks of suspicion are not too large. This can be done in many ways; for example we may require that, for a guessed π and for all $y \in \Omega_y$,

$$\begin{aligned} k_{11} &\leq P(A_1|y) \leq k_{12}, \\ (6.3) \quad k_{21} &\leq P(A_2|y) \leq k_{22}, \\ k_{31} &\leq P(A_3|y) = 1 - P(A_1|y) - P(A_2|y) \leq k_{32}. \end{aligned}$$

By analogy with the single-dimensional case, we may also require that these conditions hold for every π in some set Π believed to contain the true value. At the end of this subsection this approach will be further discussed.

Suppose that π is fixed. The relations (6.3) define a risk restriction set Q_π in which $(P(A_1|y), P(A_2|y))$ may assume values. In principle any convex set within the triangle $(0,0), (0,1), (1,0)$ could be used. However, we shall mainly consider convex polygons, since they are the results of linear restrictions. Moreover, any convex set can be regarded as the limit of such polygons with an increasing number of sides.

A plan with the risks of suspicion $(P(A_1|y), P(A_2|y))$ within a given convex set Q_π for every possible answer y and for some fixed π is said to be a Q_π -plan. When a risk restriction set Q_π is determined we wish to choose a Q_π -plan that makes $V(\lambda^*)$ small; $\lambda = \sum t_i \pi_i$. As usual in multi-parameter design problems, no plan is optimal for estimating every linear combination; however, the concept of admissible plan can be introduced:

DEFINITION A Q_π -plan is admissible in the set of all Q_π -plans if there does not exist another Q_π -plan whose Cramér-Rao bound for $\lambda = \sum t_i \pi_i$ is smaller or equally small for all t_1, t_2 and smaller for some t_1, t_2 .

The Q_π -plans are thus ordered by their corresponding information matrices (see Section 6.1). The following theorem expresses one important property of an admissible plan.

THEOREM 6a Let the risk restriction set Q_π be a convex polygon. Then an admissible Q_π -plan has the random risk of suspicion $(P(A_1|Y; \pi), P(A_2|Y; \pi))$ distributed on the vertices of Q_π .

To prove this statement we need

LEMMA 7 Suppose that the random variable $U = (U_1, U_2)^T$ takes values with probability one in the interior or on the boundary of a convex polygon $Q(U)$. Then there exists a random variable $U' = (U'_1, U'_2)^T$ with the probability mass situated in the vertices of $Q(U)$, with the same first moments as U and such that

$$V(t_1 U'_1 + t_2 U'_2) \geq V(t_1 U_1 + t_2 U_2),$$

with equality holding for every t_1, t_2 only if U takes values with probability one on the vertices of $Q(U)$.

PROOF OF LEMMA 7 We shall show that it is possible to construct a joint four-dimensional distribution for the random variables U, U' such that the two-dimensional marginal distribution of U' has the stated properties.

Let (a_i, b_i) , $i = 1, \dots, m$, be the vertices of $Q(U)$ in some coordinate

system. Since $Q(U)$ is a convex set, we may to a given point $u = (u_1, u_2)^T$ in $Q(U)$ find corresponding non-negative quantities $p_i(u)$, $i = 1, \dots, m$, continuous in u and with sum 1 such that

$$\sum p_i(u) a_i = u_1, \quad \sum p_i(u) b_i = u_2.$$

Hence, if we define the conditional distribution of U' given $U = u$ as

$$P(U'_1 = a_i, U'_2 = b_i | U = u) = p_i(u), \quad i = 1, \dots, m,$$

it follows that

$$E(U' | U = u) = u.$$

Integrating over the distribution of U , we see that U' and U have the same mean. Further, it follows that, for given t_1, t_2 ,

$$E(t_1 U'_1 + t_2 U'_2 | U = u) = t_1 u_1 + t_2 u_2.$$

Hence

$$\begin{aligned} V(t_1 U'_1 + t_2 U'_2) &= EV(t_1 U'_1 + t_2 U'_2 | u) + VE(t_1 U'_1 + t_2 U'_2 | u) = \\ &= EV(t_1 U'_1 + t_2 U'_2 | u) + V(t_1 U_1 + t_2 U_2) \geq \\ &\geq V(t_1 U_1 + t_2 U_2) \end{aligned}$$

with equality if and only if

$$t_1 U'_1 + t_2 U'_2 = t_1 U_1 + t_2 U_2$$

with probability one. □

PROOF OF THEOREM 6a Let the risk restriction set be a convex polygon Q_π .

Suppose that the response densities $f_1(y), f_2(y), f_3(y)$ fulfil the requirements of a Q_π -plan, that is $(P(A_1|Y), P(A_2|Y))$ takes for fixed π values in Q_π with probability one. Then

$$(6.4) \quad U = \begin{bmatrix} \frac{P(A_1|Y)}{\pi_1} - \frac{P(A_3|Y)}{\pi_3} \\ \frac{P(A_2|Y)}{\pi_2} - \frac{P(A_3|Y)}{\pi_3} \end{bmatrix} = \begin{bmatrix} \frac{1}{\pi_1} + \frac{1}{\pi_3} & \frac{1}{\pi_3} \\ \frac{1}{\pi_3} & \frac{1}{\pi_2} + \frac{1}{\pi_3} \end{bmatrix} \begin{bmatrix} P(A_1|Y) \\ P(A_2|Y) \end{bmatrix} - \begin{bmatrix} \frac{1}{\pi_3} \\ \frac{1}{\pi_3} \end{bmatrix}$$

assumes with probability one values in another polygon $Q(U)$ and further the expected values $E(U_1) = E(U_2) = 0$. According to Lemma 7 there exists a random variable U' with the probability mass in the vertices of $Q(U)$

and with $E(U'_1) = E(U'_2) = 0$ such that

$$(6.5) \quad V(t_1 U'_1 + t_2 U'_2) \geq V(t_1 U_1 + t_2 U_2)$$

with equality for every t_1, t_2 only if $U = U'$ with probability one.

Assume that there exists an RR-plan with answer Y' such that

$$\left[\frac{P(A_1|Y')}{\pi_1} - \frac{P(A_3|Y')}{\pi_3}, \frac{P(A_2|Y')}{\pi_2} - \frac{P(A_3|Y')}{\pi_3} \right]$$

is distributed as (U'_1, U'_2) ; we shall soon prove that this is true. Then the information matrix based on Y' has the elements

$$I'_{ij} = \text{Cov}(U'_i, U'_j), \quad i, j = 1, 2,$$

whereas the information matrix based on Y has the elements $I_{ij} = \text{Cov}(U_i, U_j)$, $i, j = 1, 2$. From (6.5) we have

$$\sum t_i t_j I'_{ij} \geq \sum t_i t_j I_{ij}$$

with equality for all t_1, t_2 only if $U = U'$ with probability one. This means that the concentration ellipse (see Cramér (1946)) for an efficient estimate of π based on Y' lies within the concentration ellipse for an efficient estimate based on Y , and the ellipses coincide only if $U = U'$ with probability one. Hence, by the definition of admissibility, the theorem is proved apart from the above assumption.

To verify the assumption concerning the existence of an RR-plan with answers Y' is equivalent to showing that there exists an RR-plan such that

$$(P(A_1|Y'), P(A_2|Y')) = (a_{1i}, a_{2i})$$

with probability p_i , $i = 1, \dots, m$, $\sum p_i = 1$, where the p_i 's are given and (a_{1i}, a_{2i}) , $i = 1, \dots, m$, are the vertices of Q_π . Let y_i , $i = 1, \dots, m$, be m different answers and define the response distributions

$$(6.6) \quad \begin{aligned} P(Y'=y_i|A_1) &= a_{1i}p_i/\pi_1, \\ P(Y'=y_i|A_2) &= a_{2i}p_i/\pi_2, \\ P(Y'=y_i|A_3) &= (1-a_{1i}-a_{2i})p_i/\pi_3, \quad i=1,2, \dots, m. \end{aligned}$$

That these distributions are probability distributions follows from the facts that $a_{ij} \geq 0$, $(1-a_{1i}-a_{2i}) \geq 0$ and $\sum_i p_i a_{ji} = \pi_j$, $i=1,2, \dots, m$, $j=1,2$. The probability of a y_i -answer is

$$P(Y'=y_i) = \sum_j P(Y'=y_i|A_j)\pi_j = p_i, \quad i=1,2, \dots, m,$$

and the risks of suspicion associated with a y_i -answer are

$$P(A_j|Y'=y_i) = a_{ji}p_i\pi_j/\pi_jp_i = a_{ji}, \quad i=1,2, \dots, m, \quad j=1,2.$$

Hence the assumption is verified. $\square$

REMARK From a result of Boes (1966) it follows that the Cramér-Rao bound is attained for small samples if and only if $m = 3$, i.e. for three-answer plans.

The restriction of $(P(A_1|y), P(A_2|y))$ to a risk restriction set Q_π for some fixed π imposes conditions on the response densities $f_i(y)$, $i = 1, 2, 3$. Typically we have, if Q_π is a polygon,

$$c_{1i}P(A_1|y) + c_{2i}P(A_2|y) \leq k_i, \quad i = 1, \dots, N,$$

for some given constants c_{ji}, k_i . This is equivalent to

$$\frac{c_{1i}\pi_1f_1(y) + c_{2i}\pi_2f_2(y)}{\pi_1f_1(y) + \pi_2f_2(y) + \pi_3f_3(y)} \leq k_i, \quad i = 1, \dots, N,$$

or

$$(6.7) \quad \pi_1(k_i - c_{1i})f_1(y) + \pi_2(k_i - c_{2i})f_2(y) + \pi_3k_if_3(y) \geq 0, \quad i = 1, \dots, N,$$

meaning that $(f_1(y), f_2(y), f_3(y))$ may only take values in a convex cone bounded by the planes (6.7) and having the vertex at the origin. The intersections of these planes form radii, which are mapped on the vertices of the risk restriction set Q_π .

By analogy with the single-proportion case one could require that $(P(A_1|y), P(A_2|y))$ should take values in a polygon Q_π for every π in some set Π believed to contain the true value. Then the restrictions (6.7) are to hold for every $\pi \in \Pi$, implying that $(f_1(y), f_2(y), f_3(y))$ should take values with probability one in a convex cone K formed by the intersections of the cones (6.7) with $\pi \in \Pi$. We shall add some comments concerning this case, without going into details.

A plan for which $(f_1(y), f_2(y), f_3(y))$ takes values only in K may be called a K -plan. For a fixed π the cone K is mapped by $(P(A_1|y), P(A_2|y))$ on a convex risk restriction set $Q(K, \pi)$, and the boundary of K is mapped

on the boundary of $Q(K, \pi)$. Hence K -plans with $(f_1(y), f_2(y), f_3(y))$ taking values with probability one on the boundary of K are preferable. This holds independently of the true value of π .

Finally we shall briefly comment upon the general case of k groups, $A_1, \dots, A_k$, with unknown proportions $\pi_1, \dots, \pi_k$, $\sum \pi_i = 1$, and response densities $f_1(y), \dots, f_k(y)$. A Q_π -plan may be defined by requiring that, for fixed $\pi = (\pi_1, \dots, \pi_{k-1})^T$ the risk of suspicion $(P(A_1|y), \dots, P(A_{k-1}|y))$ should only take values within a $(k-1)$ -dimensional simplex Q_π . Analogously to the two-dimensional case we then have the following generalization of Theorem 6a:

THEOREM 6b Let the risk restriction set Q_π be a convex $(k-1)$ -dimensional simplex. Then an admissible Q_π -plan has the random risk of suspicion $(P(A_1|Y; \pi), \dots, P(A_{k-1}|Y; \pi))$ distributed on the vertices of Q_π .

6.3. Estimation of a specified linear combination of proportions

We shall in this subsection discuss the interesting design problem that arises when one wants to estimate a single linear combination $\lambda = t_1\pi_1 + t_2\pi_2$ (assuming that there are three groups A_1, A_2, A_3 in unknown proportions π_1, π_2, π_3). As before, we consider Q_π -plans for some fixed π . The simplest design would be a two-answer plan, and we shall show that any given λ can be estimated using such a plan. However, a two-answer plan is in general not optimal, i.e. there is another Q_π -plan with smaller Cramér-Rao bound (CR-bound) for the parameter λ . Later we shall determine the optimal plan, which turns out to be a four-answer plan.

For a two-answer plan, let the response probabilities

$$\begin{aligned} p_{1i} &= P(\text{yes} | A_i), \\ p_{2i} &= P(\text{no} | A_i) = 1 - p_{1i}, \quad i = 1, 2, 3, \end{aligned}$$

be given; we shall later show how they should be chosen. Then

$$\begin{aligned} p_1 &= P(\text{yes}) = \sum_{i=1}^3 \pi_i p_{1i}, \\ p_2 &= P(\text{no}) = \sum_{i=1}^3 \pi_i p_{2i} = 1 - p_1. \end{aligned}$$

We have the risks of suspicion

$$P(A_i | \text{yes}) = \pi_i p_{1i} / p_1,$$

$$P(A_i | \text{no}) = \pi_i p_{2i} / p_2, \quad i = 1, 2,$$

and the probability of taking an A_i -individual is

$$p_1 P(A_i | \text{yes}) + p_2 P(A_i | \text{no}) = \pi_i, \quad i = 1, 2.$$

These relations yield

$$(6.8) \quad \frac{P(A_1 | \text{yes}) - \pi_1}{P(A_2 | \text{yes}) - \pi_2} = \frac{P(A_1 | \text{no}) - \pi_1}{P(A_2 | \text{no}) - \pi_2} = \frac{\pi_1 (p_{11} - p_{13})(\pi_1 - 1) + (p_{12} - p_{13})\pi_2}{\pi_2 (p_{11} - p_{13})\pi_1 + (p_{12} - p_{13})(\pi_2 - 1)}.$$

Hence the risks of suspicion $(P(A_1 | y), P(A_2 | y))$, $y = \text{yes, no}$, are situated on a straight line through the point (π_1, π_2) .

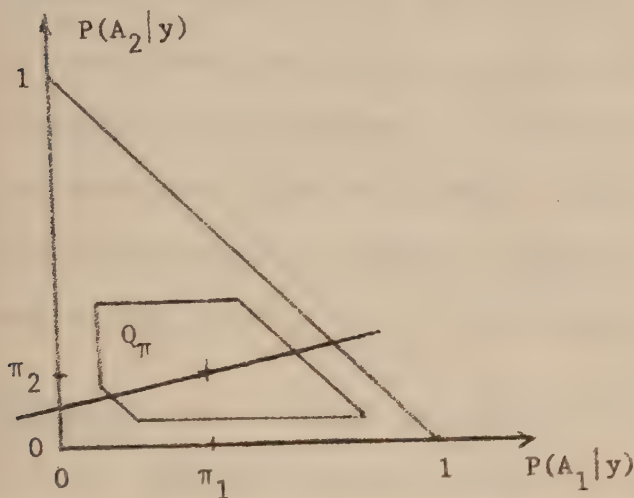


Figure 2. Risk restriction set Q_π . The risks of suspicion of a two-answer plan are situated on a straight line through (π_1, π_2) .

We shall now consider the choice of the response probabilities p_{1i} , $i = 1, 2, 3$. The probability of a yes-answer is

$$\begin{aligned} p_1 &= \pi_1 p_{11} + \pi_2 p_{12} + (1 - \pi_1 - \pi_2) p_{13} = \\ &= \pi_1 (p_{11} - p_{13}) + \pi_2 (p_{12} - p_{13}) + p_{13}. \end{aligned}$$

This quantity can be estimated without bias by the proportion of yes-answers n_1/n in a simple random sample of size n . If the p_i 's are taken so that

$$p_{11} - p_{13} = ct_1$$

and

$$p_{12} - p_{13} = ct_2,$$

then

$$(p_1 - p_{13})/c = t_1 \pi_1 + t_2 \pi_2.$$

Hence

$$\lambda^* = (n_1/n - p_{13})/c$$

is an unbiased estimate of λ . The variance of λ^* is

$$V(\lambda^*) = p_1(1-p_1)/c^2 n = (\pi_1 t_1 + \pi_2 t_2 + p_{13}/c)(1 - \pi_1 t_1 - \pi_2 t_2 - p_{13}/c)/n.$$

By inserting the expressions for $p_{11} - p_{13}$ and $p_{12} - p_{13}$ we see that the slope of the line (6.8) is determined by π_1, π_2 and the coefficients t_1, t_2 . The intersections of this line and the boundary of Q_π yield the risks of suspicion for the optimal two-answer plan (see Figure 2).

From Section 6.2 it follows that a two-answer plan is not admissible unless the above-mentioned intersections coincide with two vertices of Q_π . This means that there exists a Q_π -plan with smaller or at least as small CR-bound for the parameter λ . In the single-proportion case it was shown that two-answer plans are optimal. The following theorem states that a four-answer plan will always be sufficient when estimating a given linear combination $\lambda = t_1 \pi_1 + t_2 \pi_2$ of two proportions.

THEOREM 9a Let a convex risk restriction set Q_π be given for a fixed (π_1, π_2) . Then there exists a plan with at most four possible answers that, in the class of Q_π -plans, minimizes the Cramér-Rao bound for the variance of an unbiased estimate of any given linear combination $\lambda = t_1 \pi_1 + t_2 \pi_2$.

PROOF Without loss of generality we assume that $t_1^2 + t_2^2 = 1$. According to (6.2) the CR-bound can be written

$$C = 1/\min_{\beta} V(X_1 - \beta X_2),$$

where

$$(6.9) \quad \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = T \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}.$$

The matrix T is orthogonal with first row (t_1, t_2) (hence the second row is for example $(t_2, -t_1)$), and $(U_1, U_2)^T$ is defined by (6.4).

Suppose a convex risk restriction polygon Q_π is given for a fixed π . This polygon is mapped by the non-singular transformations (6.4) and (6.9)

into a polygon $Q(X)$, and the vertices of Q_π are mapped on the vertices of $Q(X)$. We only have to consider admissible plans, i.e. plans such that $(P(A_1|y), P(A_2|y))$ only assumes values in the vertices of Q_π or equivalently such that $(X_1, X_2)^T$ with probability one is distributed on the vertices of $Q(X)$. An admissible plan is given by the probabilities p_i , $i = 1, \dots, m$, with which $(P(A_1|Y), P(A_2|Y))$ takes the i 'th vertex of Q_π or $(X_1, X_2)^T$ takes the i 'th vertex of $Q(X)$, $i = 1, \dots, m$.

The equations $E(P(A_j|Y)) = \pi_j$, $j=1,2$, or equivalently $E(X_j) = 0$, $j = 1,2$, impose necessary restrictions on the p_i 's. On the other hand it is clear from the final step in the proof of Theorem 6a that it is possible to construct an RR-plan from the p_i 's, if they fulfil these conditions.

Let (d_{1i}, d_{2i}) , $i = 1, \dots, m$, be the vertices of $Q(X)$. Then $p_i = P(X_1 = d_{1i}, X_2 = d_{2i})$, $i = 1, \dots, m$. From the above discussion it follows that

$$1/C = \min_{\beta} V(X_1 - \beta X_2) = \min_{\beta} \sum p_i (d_{1i} - \beta d_{2i})^2$$

should be maximized with respect to $p = (p_1, \dots, p_m)^T$ when p takes values in the set determined by the equations

$$E(X_1) = \sum p_i d_{1i} = 0,$$

$$E(X_2) = \sum p_i d_{2i} = 0,$$

$$\sum p_i = 1,$$

$$p_i \geq 0, i = 1, \dots, m.$$

Straightforward regression analysis yields

$$1/C = \sum p_i (d_{1i} - \hat{\beta} d_{2i})^2,$$

where

$$\hat{\beta} = \sum p_i d_{1i} d_{2i} / \sum p_i d_{2i}^2.$$

Consider any fixed $q = (q_1, \dots, q_m)^T$ fulfilling the equations (6.10).

Let $\hat{\beta}_q = \sum q_i d_{1i} d_{2i} / \sum q_i d_{2i}^2$ be the corresponding regression coefficient and

let the class B_q consist of those p that fulfil (6.10) and the equation

$$\hat{\beta} = \sum p_i d_{1i} d_{2i} / \sum p_i d_{2i}^2 = \hat{\beta}_q.$$

For $p \in B_q$ we have

$$\sum p_i d_{1i} = \sum p_i d_{2i} = \sum p_i (d_{1i} - \hat{\beta}_q d_{2i}) d_{2i} = \sum p_i - 1 = 0,$$

$$p_i \geq 0, \quad i = 1, \dots, m.$$

Hence B_q is a convex bounded simplex lying in the intersection of four hyperplanes in m -dimensional Euclidean space. In this simplex

$$g(p) = \min_{\beta} V(X_1 - \beta X_2) = \sum p_i (d_{1i} - \hat{\beta}_q d_{2i})^2$$

is a linear function of the p_i 's, which implies that it assumes its maximum value in one of the vertices of B_q . But in a vertex of B_q at least $m-4$ of the hyperplanes $p_i = 0$ intersect. For any q not having at least $m-4$ of the components q_i equal to zero, we can thus find a p , fulfilling (6.10), with at least $m-4$ components p_i equal to zero and with a larger or equally large value of $1/C$. Hence the theorem holds when the risk restriction set Q_π is a polygon. But any risk set can be considered as the limit of a sequence of convex polygons with an increasing number of sides, and thus the theorem holds for any convex risk set Q_π . $\square$

The above result can be generalized in a straightforward manner to a population of k groups.

THEOREM 9b Let the population consist of the classes A_i , $i = 1, \dots, k$, and let a convex risk restriction set Q_π be given for a fixed $\pi = (\pi_1, \dots, \pi_{k-1})$. Then there exists a plan with at most $2k-2$ possible answers that, in the class of Q_π -plans, minimizes the Cramér-Rao bound for the variance of an unbiased estimate of any given linear combination

$$\lambda = \sum_1^{k-1} t_i \pi_i.$$

REMARK The results of this section have a wider scope than is at first anticipated. Suppose one studies a random variable Z with arbitrary distribution. For practical purposes the sample space may then be divided into classes A_i , $i = 1, \dots, k$, and Z may be approximated with constants t_i in these classes. Different population parameters can now be expressed as linear functions of the proportions $\pi_i = P(Z \in A_i)$, $i = 1, \dots, k$, e.g. the expected value $E(Z) \approx \sum t_i \pi_i$, and the above results are applicable.

7. θ -DEPENDENT PROPORTIONS

The proportions π_i of individuals belonging to the classes A_i , $i = 1, \dots, k$, sometimes depend on a parameter $\theta = (\theta_1, \dots, \theta_r) \in \Theta$. This occurs e.g.

when a distribution $G(x; \theta)$ is partitioned into classes A_i with

$$\pi_i(\theta) = \int_{A_i} dG(x; \theta).$$

Let, as before, $f_i(y)$ be the known response density of individuals from the class A_i , $i = 1, \dots, k$. The observable answer Y then has the density

$$f_Y(y; \theta) = \pi_1(\theta)f_1(y) + \dots + \pi_k(\theta)f_k(y).$$

Assume that the derivatives

$$\frac{\partial}{\partial \theta_i} \pi_j(\theta), \quad i = 1, \dots, r, \quad j = 1, \dots, k,$$

exist for $\theta \in \Theta$ and let

$$d_{ij} = \frac{\partial}{\partial \theta_i} (\pi_j(\theta)) / \pi_j(\theta) = \frac{\partial}{\partial \theta_i} \log \pi_j(\theta), \quad i = 1, \dots, r, \\ j = 1, \dots, k,$$

Then the elements of the information matrix I are

$$I_{ij} = E \left\{ \sum_v \frac{\partial}{\partial \theta_i} (\pi_v(\theta) f_v(Y)) / f_Y(Y; \theta) \cdot \sum_v \frac{\partial}{\partial \theta_j} (\pi_v(\theta) f_v(Y)) / f_Y(Y; \theta) \right\} = \\ = (d_{i1}, \dots, d_{ik}) E \begin{bmatrix} P(A_1|Y) \\ \vdots \\ P(A_k|Y) \end{bmatrix} (P(A_1|Y), \dots, P(A_k|Y)) \begin{bmatrix} d_{j1} \\ \vdots \\ d_{jk} \end{bmatrix};$$

note that

$$f_v(y) \pi_v(\theta) = P(A_v|y) f_Y(y; \theta).$$

Hence the information matrix I can be written

$$I = DAD^T,$$

where $D = (d_{ij})$ is an $r \times k$ matrix and $A = (a_{ij})$ is a $k \times k$ matrix having elements

$$a_{ij} = E(P(A_i|Y)P(A_j|Y)), \quad i, j = 1, \dots, k.$$

The information matrix is thus separated into a matrix A depending on the response densities $f_v(y)$ and a matrix D formed by the derivatives of $\log \pi_i(\theta)$ and independent of $f_v(y)$.

A necessary condition for us to be able to estimate θ is that I is non-singular.

THEOREM 10 The information matrix I is singular if and only if either rank D is less than r (see p.33) or, with probability one, $(P(A_1|Y), \dots, P(A_k|Y))$ is orthogonal to some vector $x \neq 0$ spanned by the rows of D .

PROOF The information matrix I is singular if and only if there exists an $x = (x_1, \dots, x_r)^T \neq 0$, such that $x^T I x = 0$, that is if and only if

$$\begin{aligned} E(x^T D \begin{bmatrix} P(A_1|Y) \\ \vdots \\ P(A_k|Y) \end{bmatrix} (P(A_1|Y), \dots, P(A_k|Y)) D^T x) = \\ = \int (x^T D \begin{bmatrix} P(A_1|Y) \\ \vdots \\ P(A_k|Y) \end{bmatrix})^2 f_Y(y; \theta) d\mu(y) = 0 \end{aligned}$$

for some $x \neq 0$. This is equivalent to

$$x^T D \begin{bmatrix} P(A_1|Y) \\ \vdots \\ P(A_k|Y) \end{bmatrix} = 0$$

with probability one for some $x \neq 0$. Hence the theorem is proved. $\square$

REMARK Evidently rank D is less than r if $r > k$. Moreover, this holds when $r = k$ since, in view of the fact that $\sum_{i=1}^k \pi_i(\theta) = 1$, one column of D can then be written as a linear combination of the others.

Now let θ be a one-dimensional parameter. Then Fisher's information is

$$I = E(\sum_{i=1}^k d_{\nu} P(A_{\nu}|Y))^2,$$

where $d_{\nu} = \frac{d}{d\theta} \log \pi_{\nu}(\theta)$. Suppose that a convex risk restriction polygon Q_{π} is given, in which $P(A|Y) = (P(A_1|Y), \dots, P(A_k|Y))^T$ is to be distributed for some fixed $\theta = \theta_0$. We only consider plans such that $P(A|Y)$ with probability one takes values in the vertices of Q_{π} , $a_i = (a_{1i}, \dots, a_{ki})^T$, $i = 1, \dots, m$. Let p_i be the probability that $P(A|Y)$ assumes the vertex

a_i , $i = 1, \dots, m$, and let $p = (p_1, \dots, p_m)^T$. Then we have

$$(7.1) \quad \begin{aligned} \sum_i a_{iv} p_i &= \pi_v(\theta_0), \quad v = 1, \dots, k, \\ p_i &\geq 0, \quad i = 1, \dots, m, \end{aligned}$$

and, moreover, a plan is determined by p if p fulfils (7.1) (see end of the proof of Theorem 6a).

We wish to find a plan that maximizes Fisher's information I about θ . The equations (7.1) define a convex simplex B_p with at least $m-k$ dimensions, in which p may take values. Moreover, Fisher's information

$$I = \sum_i p_i \left(\sum_v d_v a_{iv} \right)^2$$

is a linear function of p , and so maximizing I is an ordinary linear-programming problem. The maximum is attained for p in a vertex of B_p , where at least $m-k$ of the hyperplanes $p_i = 0$, $i = 1, \dots, m$, intersect. Hence the solution has at most k of the p_i 's larger than zero meaning that a k -answer plan is always sufficient. This can be compared with the multi-proportion case of the last section, where an optimal plan generally has $2k-2$ possible answers and besides is numerically more difficult to determine.

EXAMPLE 2 Let the sensitive variable X be $N(\theta, 1)$ and partition the sample space of X into the classes $A_1 = (-\infty, -1)$, $A_2 = (-1, 1)$ and $A_3 = (1, \infty)$. Then

$$\pi_1(\theta) = \Phi(-1-\theta), \quad \pi_2(\theta) = \Phi(1-\theta) - \Phi(-1-\theta), \quad \pi_3(\theta) = 1 - \Phi(1-\theta),$$

$$\pi_1'(\theta) = -\phi(-1-\theta), \quad \pi_2'(\theta) = -\phi(1-\theta) + \phi(-1-\theta), \quad \pi_3'(\theta) = \phi(1-\theta),$$

where Φ and ϕ are the distribution function and the density of a standard normal variable. Now assume that $\theta = \theta_0 = 0$. Then

$$\pi_1(0) = \pi_3(0) = \Phi(-1) = 0.1587, \quad \pi_2(0) = \Phi(1) - \Phi(-1) = 0.6826,$$

$$\pi_1'(0) = -1/\sqrt{2\pi e} = -0.2420, \quad \pi_2'(0) = 0, \quad \pi_3'(0) = 1/\sqrt{2\pi e} = 0.2420.$$

Hence Fisher's information is

$$\begin{aligned} I &= E\{\Sigma P(A_v|Y) \pi_v'(0) / \pi_v(0)\}^2 = \\ &= E\{P(A_1|Y) - P(A_3|Y)\}^2 / 2\pi e (\Phi(-1))^2. \end{aligned}$$

Suppose that A_1 and A_3 are stigmatizing characteristics. Then the risk restriction set Q_π for $\theta = 0$ could be chosen as in Figure 3.

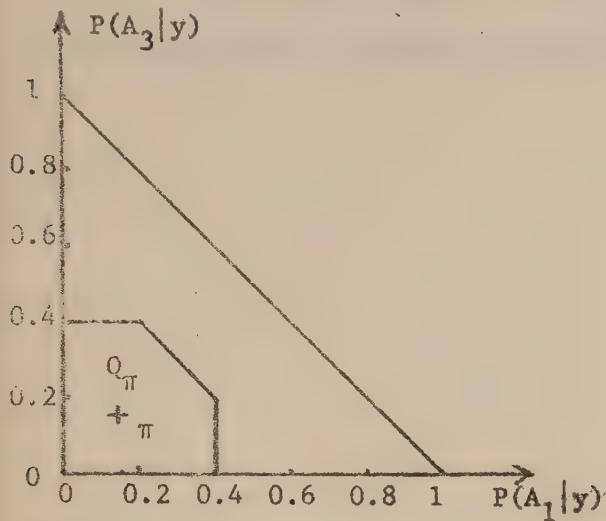


Figure 3. Risk restriction set Q_π when the characteristics A_1 and A_3 are supposed to be stigmatizing.

Let p_i , $i = 1, \dots, 5$, be the probabilities that $(P(A_1|Y), P(A_3|Y))$ assumes the vertices $(0,0)$, $(.4,0)$, $(.4,.2)$, $(.2,.4)$ and $(0,.4)$.

Then $p = (p_1, \dots, p_5)$ must fulfil the following conditions

$$0.2p_4 + 0.4(p_2 + p_3) = 0.1587,$$

$$0.2p_3 + 0.4(p_4 + p_5) = 0.1587,$$

$$p_1 + \dots + p_5 = 1,$$

$$p_i \geq 0, \quad i = 1, \dots, 5.$$

The information

$$\begin{aligned} I &= \frac{1}{2\pi e(\phi(-1))^2} (0.4^2 p_2 + 0.2^2 (p_3 + p_4) + 0.4^2 p_5) = \\ &= \text{constant} \cdot (4(p_2 + p_5) + (p_3 + p_4)) \end{aligned}$$

is to be maximized under the conditions above. By symmetry I is maximized

for $p_2 = p_5$ and $p_3 = p_4$. The equations yield

$$p_2 = 2.5(0.1587 - 0.6p_3) \geq 0,$$

$$p_1 = p_3 + 0.2065 \geq 0.$$

Hence

$$0 \leq p_3 \leq 0.2065.$$

But $I = \text{constant} \cdot (1.5870 - 5p_3)$ is then maximized for $p_3 = 0$, which

makes $p_4 = 0$, $p_2 = p_5 = 0.3968$ and $p_1 = 0.2065$. The maximal information, attained with these p_i 's, is $I = 0.2951 = 1/3.39$. This value should be compared with the information from a direct observation of X , which is 1.

REFERENCES

- Abul-Ela, A.-L.A., Greenberg, B.G., Horvitz, D.G. (1967). A multi-proportions randomized response model. *J. Amer. Statist. Assoc.* 62, 990-1008.
- Boes, D.C. (1966). On the estimation of mixing distributions. *Ann. Math. Statist.* 37, 177-188.
- Boruch, R.F. (1972). Relations among statistical methods for assuring confidentiality of social research data. *Social Science Research* 1, 403-414.
- Bourke, P.D., Dalenius T. (1973). Multi-proportions randomized response using a single sample. Report no. 68 of the research project Errors in Surveys, Institute of Statistics, University of Stockholm.
- Bourke, P.D. (1974). Symmetry of response in randomized response designs. Report no. 75 of the research project Errors in Surveys, Institute of Statistics, University of Stockholm.
- Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.
- Eriksson, S. (1973). Randomized Interviews for Sensitive Questions. Ph. D. Thesis, Institute of Statistics, University of Göteborg.
- Greenberg, B.G., Abul-Ela, A.-L.A., Simmons, W.R., Horvitz, D.G. (1969). The unrelated question randomized response model: Theoretical framework. *J. Amer. Statist. Assoc.* 64, 520-539.
- Greenberg, B.G., Kuebler Jr, R.R., Abernathy, J.R., Horvitz, D.G. (1971). Application of the randomized response technique in obtaining quantitative data. *J. Amer. Statist. Assoc.* 66, 243-250.
- Lanke, J. (1975a). On the choice of the unrelated question in Simmons' version of randomized response. *J. Amer. Statist. Assoc.* 70, 80-83.
- Lanke, J. (1975b). On the degree of protection in randomized interviews. Report no. 2 of the research project Confidentiality in Surveys, Institute of Statistics, University of Stockholm. (Invited paper for the IASS session in Warsaw, September 1975).
- Moors, J.J.A. (1971). Optimization of the unrelated question randomized response model. *J. Amer. Statist. Assoc.* 66, 627-629.
- Teicher, H. (1963). Identifiability of finite mixtures. *Ann. Math. Statist.* 34, 1265-1269.
- Warner, S.L. (1965). Randomized response: A technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.* 60, 63-69.
- Warner, S.L. (1971). The linear randomized response model. *J. Amer. Statist. Assoc.* 66, 884-888.

Paper [B]

EFFICIENCY VERSUS PROTECTION IN A GENERAL RANDOMIZED RESPONSE MODEL

Department of Mathematical Statistics

Box 725

S-220 07 LUND, Sweden

EFFICIENCY VERSUS PROTECTION IN A
GENERAL RANDOMIZED RESPONSE MODEL

by

Harald Anderson

No. 1975:10

July 1975

EFFICIENCY VERSUS PROTECTION IN A GENERAL RANDOMIZED RESPONSE MODEL

by

Harald Anderson

1. INTRODUCTION AND SUMMARY

The randomized response interviewing technique (RR for short) was introduced by Warner in 1965 as a means of reducing answer bias and answer refusal when dealing with embarrassing subjects. Warner originally considered the case of estimating the proportion of individuals in a population who have some stigmatizing characteristic, e.g. that of being a tax evader. Since 1965 several RR-models have been proposed, for estimating both a single proportion (Greenberg et al. (1969)), several proportions (Abul-Ela et al. (1967)) and other population parameters such as the mean of some quantitative characteristic (Greenberg et al. (1971), Eriksson (1973), Poole (1974), Dalenius et al. (1974)).

The core of RR is a randomizing device, which gives answers partly random and partly dependent on the value of the respondent's sensitive characteristic. The idea is that, from these answers, one should be able to obtain estimates of the population parameters without direct knowledge of the sensitive characteristic of the sampled individuals.

Different randomizing devices give various degrees of protection to the interviewee's privacy and also allow estimates of different efficiencies. The aim of an RR-plan is therefore twofold:

- (i) to afford good estimates of the population parameters,
- (ii) to attain a sufficient degree of protection to the interviewee.

Most research work on RR has dealt with questions concerning construction and test of different RR-schemes in order to decrease both the answer bias and the answer refusal in special situations. Then the design of the randomizing mechanism and the behaviour of the interviewer play important roles. This approach to RR is very natural and essential from a practical point of view.

In this paper are considered some theoretical aspects of randomized response, which make it possible to insert the RR-technique in a wider statistical context. Some work in this direction has been done by Warner (1971), who introduced a "linear RR-model". He assumes that the embarrassing variable X is p -dimensional, and that one wants to estimate some linear function of the mean vector $E(X)$. In a randomizing mechanism $r \times p$ matrices T are generated according to a known distribution independently of X , and a respondent with the true value $X = x$ is asked to answer $Y = Tx$. This model covers some of the above-mentioned situations.

Section 2 of this paper introduces a general RR-model, where the answer from an individual with the characteristic $X = x$ follows a distribution depending on x . The family of such response distributions and the population distribution of X are the essential ingredients of an RR-scheme from a statistical point of view. It will be seen, for example, that the observable answer from a randomly taken individual is distributed according to a mixture of the family of response distributions with the population distribution as a mixing distribution.

In Section 3 it is studied how the statistical inference about the population is affected by introducing RR. If the distribution of X is regular, the loss of efficiency when estimating a parameter is assessed by means of Fisher's conditional information. The Kullback-Leibler information is used to estimate the loss of power due to RR when testing two simple hypotheses. In Section 4 the privacy aspects of RR are dealt with starting from the family of conditional distributions of X given the different possible RR-answers. Various measures of the protection provided by an RR-scheme are considered and some planning ideas are proposed.

In Section 5 two examples are given. In the first example the density of X is assumed to belong to an exponential family and in the second example a special RR-scheme, known as the "model with given answers" (Eriksson (1973)), is considered. Both examples illustrate how increased protection causes loss of information about the population.

2. A GENERAL RANDOMIZED RESPONSE MODEL

Suppose that we have a population of individuals with some embarrassing characteristic $X \in \Omega_X$ distributed according to the unknown distribution $F_X(x)$. Our aim is to estimate one or several population parameters, e.g. the mean of X , $E(X)$, or some proportion $F_X(c)$, and for this purpose we shall take a random sample of persons to be interviewed. However, for privacy reasons the sampled individuals may not be interviewed directly. Thus we have to ask questions in such a way that, from the answers, the population parameters can be estimated but the values of the respondents cannot be inferred. This can be attained by introducing a random device known to the statistician and producing an answer Y with a probability distribution dependent on the X -value of the respondent. More precisely, we have the general RR-model:

An individual with $X = x$ answers according to the known probability density $h_Y(y|x)$, $x \in \Omega_X$.

The densities $h_Y(y|x)$, $x \in \Omega_X$, are called response densities and they hold with respect to some measure $d\lambda(y)$ thus including continuous as well as discrete distributions. In general the family of response densities $\{h_Y(y|x); x \in \Omega_X\}$ is chosen by the statistician. Nearly any family can be used by choosing an appropriate randomizing device.

The statistician can only observe the answer Y from a randomly taken individual. It has the density

$$f_Y(y) = \int h_Y(y|x) dF_X(x) = E(h_Y(y|X)),$$

a mixture of the different response densities with $F_X(x)$ as mixing distribution. The population parameters of X may be inferred from observations of Y , whereas the single X -values may not be inferred, if the densities $h_Y(y|x)$, $x \in \Omega_X$, overlap.

Most RR-models proposed so far have been analyzed in terms of their randomizing mechanisms. However, as will be shown by examples at the end of this section and in the following sections, deeper insight into the nature of such models is gained by introducing the relevant family of response

distributions. Indeed, this family determines the two most important characteristics of the RR-scheme, namely

- (i) the information provided about the population,
- (ii) the degree of protection of the interviewee's privacy.

These two aspects, which will play an important role in the sequel, are fundamentally conflicting; a large protection to the respondent is incompatible with large information about the population.

To study (i) and (ii) in detail we shall need the family of conditional densities of X given $Y = y$, $y \in \Omega_y$:

$$(2.1) \quad h_X(x|y) = h_Y(y|x)f_X(x)/f_Y(y), \quad y \in \Omega_y,$$

Ω_y being the set of possible RR-answers. In the following this family will be called the family of revealing densities. It depends both on the response distributions and on the unknown population distribution. The importance of the revealing densities for (i) will appear from the next section. That they are important also for (ii) is obvious. See further Section 4.

EXAMPLE 1 Suppose we want to estimate the proportion θ of tax evaders in a population from a simple random sample of n individuals taken with replacement. Let the random variable X take the value one for a tax evader and zero otherwise, and let the interviewees answer according to the response densities $h_Y(y|1)$ and $h_Y(y|0)$, respectively. This can be arranged in practice by using a pack of cards with two answers on each card or by a spinner with two scales, one for each category. The observable answer Y then has the density

$$f_Y(y;\theta) = \theta h_Y(y|1) + (1-\theta)h_Y(y|0),$$

a mixture between two known densities $h_Y(y|1)$ and $h_Y(y|0)$.

The probability of being a tax evader given the answer $Y = y$ is

$$h_X(1|y;\theta) = P(X = 1|Y = y;\theta) = \frac{\theta h_Y(y|1)}{\theta h_Y(y|1) + (1-\theta)h_Y(y|0)}.$$

As can easily be verified, this example includes the well-known Warner and Simmons RR-schemes. For further discussion, see Anderson (1975a).

EXAMPLE 2 Assume that the sensitive variable X is normally distributed with unknown mean θ and variance one, i.e. $X \in N(\theta, 1)$. The interviewees are asked

to add a $N(0, \sigma)$ -distributed random number Z to their X -values. Then the answer of an individual, given his true value $X = x$, is normally distributed with mean x and variance σ^2 , $N(x, \sigma)$. The observable answer from a randomly taken individual $Y = X + Z \in N(\theta, \sqrt{1 + \sigma^2})$, and the conditional density of X given $Y = y$ is

$$\begin{aligned} h_X(x|y; \theta, \sigma) &= \frac{\frac{1}{2\pi\sigma} \exp\{-0.5[(\frac{y-x}{\sigma})^2 + (x-\theta)^2]\}}{\frac{1}{\sqrt{2\pi} \sqrt{1+\sigma^2}} \exp\{-0.5(y-\theta)^2/(1+\sigma^2)\}} = \dots \\ &= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{1+\sigma^2}{\sigma^2}} \exp\left\{-\frac{(1+\sigma^2)}{2\sigma^2} \left(x - \frac{y+\theta\sigma^2}{1+\sigma^2}\right)^2\right\}, \end{aligned}$$

that is, X given $Y = y$ belongs to $N\left(\frac{y+\theta\sigma^2}{1+\sigma^2}, \sqrt{\frac{\sigma^2}{1+\sigma^2}}\right)$. Not surprisingly, we see that the larger σ is, the more out-spread the revealing distributions are and the more the interviewees are protected.

3. THE AMOUNT OF INFORMATION PROVIDED BY RANDOMIZED RESPONSES

3.1. One-dimensional parameter

Assume that f_X belongs to a regular family of densities with respect to some measure $d\mu(x)$, $\{f_X(x;\theta); \theta \in \Theta\}$; θ is an unknown one-dimensional parameter. Fisher's information, if X is observed, is then

$$I(X) = E \left[\frac{\partial \log f_X(X;\theta)}{\partial \theta} \right]^2.$$

The amount of information provided by a sample of n independent observations is $nI(X)$. In sampling surveys the sample size is often large and then the minimum variance bound $(nI(X))^{-1}$ of the Cramér-Rao inequality is approximately attained for the maximum-likelihood estimate. Thus Fisher's information is a relevant measure of the information about θ contained in a sample.

We shall now examine how the information about θ is modified when observing the randomized response Y instead of X . Assume that we have a family of response densities $\{h_Y(y|x); x \in \Omega_X\}$ with respect to some measure $d\lambda(y)$. The observable answer from a randomly taken individual has the density

$$f_Y(y;\theta) = \int h_Y(y|x) f_X(x;\theta) d\mu(x)$$

with respect to $d\lambda(y)$. The family $\{f_Y(y;\theta); \theta \in \Theta\}$ is regular under weak conditions on the family of response densities, and in the sequel we shall assume that these conditions are fulfilled.

Consider now the random variable (X,Y) having the density

$$f_{X,Y}(x,y;\theta) = h_Y(y|x) f_X(x;\theta) = h_X(x|y;\theta) f_Y(y;\theta).$$

The total information yielded by (X,Y) is

$$I(X,Y) = E \left[\frac{\partial \log f_{X,Y}(X,Y;\theta)}{\partial \theta} \right]^2.$$

This quantity can be separated into components in two different ways (see Papaioannou and Kempthorne (1971) or Fraser (1965)):

$$(3.1) \quad I(X,Y) = I(Y) + I(X|Y) = I(X) + I(Y|X),$$

where

$$\begin{aligned} I(X|Y) &= \int f_Y(y;\theta) \int \left(\frac{\partial}{\partial \theta} \log h_X(x|y;\theta) \right)^2 h_X(x|y;\theta) d\mu(x) d\lambda(y) = \\ &= E \left[\frac{\partial}{\partial \theta} \log h_X(X|Y;\theta) \right]^2, \end{aligned}$$

and $I(Y|X)$ is defined by interchanging x and y in the above expression.

The quantity $I(X|Y)$ is thus the expected conditional information about θ from X given Y . In our case X is sufficient and hence

$$(3.2) \quad I(X,Y) = I(X).$$

Combining (3.1) and (3.2) we have

THEOREM 1 The decrease of Fisher's information due to randomized response is

$$I(X) - I(Y) = I(X|Y),$$

where $I(X|Y)$ is Fisher's conditional information from X given Y .

Suppose $\hat{\theta}(X)$ and $\hat{\theta}(Y)$ are unbiased efficient estimates of θ based on the same number of observations of X and Y , respectively. Then the variances of $\hat{\theta}(X)$ and $\hat{\theta}(Y)$ attain the Cramér-Rao bound and from Theorem 1 we have the simple relations

$$\frac{V(\hat{\theta}(Y))}{V(\hat{\theta}(X))} = 1 + \frac{I(X|Y)}{I(Y)} = \left[1 - \frac{I(X|Y)}{I(X)} \right]^{-1}.$$

The loss of information $I(X|Y)$ can intuitively be comprehended in the following way. A given answer $Y=y$ is informative about the true value of X if the conditional density $h_X(x|y;\theta)$ of X given $Y=y$ is concentrated around x regardless of the value of θ . If this is the case the derivative $\frac{\partial}{\partial \theta} \log h_X(x|y;\theta)$ tends to be small and if this holds for most of the possible answers the loss of information about θ , $I(X|Y) = E \left[\frac{\partial}{\partial \theta} \log h_X(X|Y;\theta) \right]^2$, is likewise small.

EXAMPLE 1 (continued) Estimation of a proportion. The revealing probability given the answer $Y=y$ is (in Anderson (1975a) called "risk of suspicion")

$$P(X=1|Y=y;\theta) = \frac{\theta h_Y(y|1)}{\theta h_Y(y|1) + (1-\theta) h_Y(y|0)}.$$

Put

$$p(\theta, y) = P(X=1|Y=y; \theta) = 1 - P(X=0|Y=y; \theta).$$

Then

$$I(X|Y) = \int I(X|Y=y) f_Y(y; \theta) d\lambda(y),$$

where

$$\begin{aligned} I(X|Y=y) &= \left[\frac{\partial p(\theta, y)}{\partial \theta} \right]^2 / p(\theta, y) + \left[\frac{\partial p(\theta, y)}{\partial \theta} \right]^2 / (1-p(\theta, y)) = \dots \\ &= \frac{h_Y(y|1)h_Y(y|0)}{\theta(1-\theta)[\theta h_Y(y|1) + (1-\theta)h_Y(y|0)]^2}. \end{aligned}$$

Hence we have the loss of information

$$I(X|Y) = \int \frac{h_Y(y|1)h_Y(y|0)}{\theta(1-\theta)(\theta h_Y(y|1) + (1-\theta)h_Y(y|0))} d\lambda(y).$$

If the response densities $h_Y(y|0)$ and $h_Y(y|1)$ are much alike, the loss of information is large. Then the revealing probability $p(\theta, y)$ also differs little from the apriori probability θ .

EXAMPLE 2 (continued) Estimation of the mean of a normal distribution.

The revealing distribution given the answer $Y=y$ is $N\left(\frac{y+\theta\sigma^2}{1+\sigma^2}, \sqrt{\sigma^2/(1+\sigma^2)}\right)$, and since the ordinary maximum-likelihood estimate of θ is then efficient, we have

$$I(X|Y=y) = \left[\left(\frac{1+\sigma^2}{\sigma^2} \right)^2 V(X|Y=y) \right]^{-1} = \sigma^2/(1+\sigma^2).$$

Hence the loss of information due to RR is

$$I(X|Y) = \sigma^2/(1+\sigma^2),$$

an increasing function of σ^2 .

3.2. Multi-dimensional parameter

The results for a single-dimensional parameter in a natural way generalize to an r -dimensional parameter θ ($r > 1$). Assume that $\{f_X(x; \theta); \theta \in \Theta\}$ is a regular family of densities with the parameter θ r -dimensional. Assume further that the family of response densities $\{h_Y(y|x); x \in \Omega_x\}$ fulfils necessary regularity conditions. The information matrix $I(X)$ from a direct observation of X has the elements

$$I_{ij}(X) = E\left[\frac{\partial}{\partial\theta_i} \log f_X(X;\theta) \cdot \frac{\partial}{\partial\theta_j} \log f_X(X;\theta)\right], \quad i,j = 1, \dots, r,$$

and the information matrix $I(Y)$ from a randomized response Y has the elements

$$I_{ij}(Y) = E\left[\frac{\partial}{\partial\theta_i} \log f_Y(Y;\theta) \cdot \frac{\partial}{\partial\theta_j} \log f_Y(Y;\theta)\right], \quad i,j = 1, \dots, r.$$

It can be shown that relation (3.1) also holds for information matrices (see Papaioannou and Kempthorne (1971)). The elements of $I(X|Y)$ are

$$I_{ij}(X|Y) = \int f_Y(y;\theta) \int \frac{\partial}{\partial\theta_i} \log h_X(x|y;\theta) \cdot \frac{\partial}{\partial\theta_j} \log h_X(x|y;\theta) h_X(x|y;\theta) d\mu(x) d\lambda(y),$$

$$i,j = 1, \dots, r,$$

that is, $I(X|Y)$ is the information matrix from the random variable X given Y . Since $I(X|Y)$ is positive semidefinite and X is sufficient, we have that Theorem 1 also holds for a multi-dimensional parameter.

EXAMPLE 3 Let the population consist of three classes characterized by

$X = 1, 2, 3$ and let $\theta_i = P(X=i)$, $i = 1, 2$, be the unknown parameters.

Then $P(X=3) = 1 - \theta_1 - \theta_2$. Individuals characterized by $X=i$ answer according to the density $h_Y(y|i)$, $i = 1, 2, 3$. The answer from a randomly chosen individual then has the density

$$f_Y(y; \theta_1, \theta_2) = \theta_1 h_Y(y|1) + \theta_2 h_Y(y|2) + (1 - \theta_1 - \theta_2) h_Y(y|3).$$

The revealing probabilities are

$$P(X=i|Y=y; \theta_1, \theta_2) = \theta_i h_Y(y|i) / f_Y(y; \theta_1, \theta_2), \quad i = 1, 2,$$

and

$$P(X=3|Y=y; \theta_1, \theta_2) = (1 - \theta_1 - \theta_2) h_Y(y|3) / f_Y(y; \theta_1, \theta_2)$$

for an interviewee who has answered $Y=y$. In this case the information matrix based on Y can easily be calculated. It has the elements

$$I_{ij}(Y) = E\left[\frac{\partial \log f_Y(Y; \theta_1, \theta_2)}{\partial \theta_i} \cdot \frac{\partial \log f_Y(Y; \theta_1, \theta_2)}{\partial \theta_j}\right] =$$

$$= E\left[\frac{h_Y(Y|i) - h_Y(Y|3)}{f_Y(Y; \theta_1, \theta_2)} \cdot \frac{h_Y(Y|j) - h_Y(Y|3)}{f_Y(Y; \theta_1, \theta_2)}\right], \quad i,j = 1, 2.$$

If the response densities are much alike, the elements of $I(Y)$ tend to be very small. On the other hand, the revealing probabilities then differ little from the apriori values.

3.3. Testing of two simple hypotheses

The research on RR reported so far in the literature only concerns estimation problems. We shall now consider the problem of testing two simple hypotheses.

The embarrassing variable X is assumed to be distributed according to either of the hypotheses:

$$H_0: f_X^0(x)$$

$$H_1: f_X^1(x).$$

In the theory of hypothesis testing the Kullback-Leibler discriminating information of $f_X^1(x)$ in favour of $f_X^0(x)$,

$$I(1;0;X) = \int f_X^1(x) \log \frac{f_X^1(x)}{f_X^0(x)} d\mu(x),$$

may be regarded as the counterpart to Fisher's information in the theory of estimation. If the significance level α of a likelihood-ratio test for testing $f_X^1(x)$ against $f_X^0(x)$ is fixed and β_n is the probability of a type two error when the number of observations is n , then (Kullback (1959), p.77)

$$\lim_{n \rightarrow \infty} \beta_n^{1/n} = \exp(-I(1;0;X)).$$

Suppose as before that an x -individual answers according to the density $h_Y(y|x)$. Then the observable answer from a randomly taken individual has the density

$$f_Y^0(y) = \int h_Y(y|x) f_X^0(x) d\mu(x)$$

or

$$f_Y^1(y) = \int h_Y(y|x) f_X^1(x) d\mu(x)$$

according to whether H_0 or H_1 is true. The Kullback-Leibler information based on Y is

$$I(1;0;Y) = \int f_Y^1(y) \log \frac{f_Y^1(y)}{f_Y^0(y)} d\lambda(y).$$

Now consider the random variable (X,Y) . A general information relation says (see Kullback (1959))

$$I(1;0;X,Y) = I(1;0;Y) + I(1;0;X|Y) = I(1;0;X) + I(1;0;Y|X),$$

where

$$I(1;0;X|Y) = \int f_Y^1(y) \int h_X^1(x|y) \log \frac{h_X^1(x|y)}{h_X^0(x|y)} d\mu(x) d\lambda(y)$$

with

$$h_X^i(x|y) = h_Y(y|x) f_X^i(x) / f_Y^i(y), \quad i = 0, 1,$$

and $I(1:0; Y|X)$ is defined analogously. Since X is sufficient for (X, Y)

we have (see Kullback (1959)) $I(1:0; X, Y) = I(1:0; X)$ and hence

THEOREM 2 The decrease of Kullback-Leibler's discriminating information due to randomized response is

$$I(1:0; X) - I(1:0; Y) = I(1:0; X|Y),$$

where $I(1:0; X|Y)$ is Kullback-Leibler's conditional information from X given Y .

EXAMPLE 2 (continued) Assume that $X \in N(0, 1)$ under H_0 and that $X \in N(\theta, 1)$ under H_1 . After adding $Z \in N(0, \sigma^2)$ the observable variable $Y = X + Z$ has the distribution $N(0, \sqrt{1 + \sigma^2})$ under H_0 and $N(\theta, \sqrt{1 + \sigma^2})$ under H_1 . The revealing distributions given the answer $Y = y$ are $N(y/(1 + \sigma^2), \sqrt{\sigma^2/(1 + \sigma^2)})$ and $N((y + \theta\sigma^2)/(1 + \sigma^2), \sqrt{\sigma^2/(1 + \sigma^2)})$, respectively.

We immediately have

$$I(1:0; X) = E_1 \left(\log \frac{f_X^1(X)}{f_X^0(X)} \right) = 0.5 E_1 (-(X - \theta)^2 + X^2) = 0.5 \theta^2,$$

where E_1 denotes expectation with respect to $f_X^1(x)$. Analogously the information from Y is $I(1:0; Y) = \theta^2 / 2(1 + \sigma^2)$ and hence the conditional information is

$$I(1:0; X|Y) = I(1:0; X) - I(1:0; Y) = \theta^2 \sigma^2 / 2(1 + \sigma^2).$$

It is seen that large values of σ^2 increase the loss of information $I(1:0; X|Y)$.

4. PRIVACY ASPECTS

The choice of the family of response distributions determines the degree of protection of the interviewee's privacy as well as the possibility of obtaining good estimates of the population parameters. In Section 3 we saw how the loss of Fisher's information (and Kullback-Leibler's information) depends on the family of revealing densities through the conditional information $I(X|Y)$. We shall now consider the protection of the sampled individual's privacy. Some measures of the protection provided by an RR-scheme will be introduced, and then two qualities are desirable: the measure should, of course, be relevant from the point of view of RR and it should be in harmony with statistical theory.

Let us begin with a general discussion. Before the interview the density $f_X(x;\theta)$ provides knowledge about the variable X of a randomly sampled individual. Suppose that he answers $Y=y$. Then the revealing density $h_X(x|y;\theta)$ yields information about X and the discrepancy between $f_X(x;\theta)$ and $h_X(x|y;\theta)$ reflects the invasion of privacy caused by the answer $Y=y$. If $h_X(x|y;\theta)$ is very concentrated around some value x , the protection provided by the answer $Y=y$ is very small; conversely, if $h_X(x|y;\theta)$ is widely spread, privacy is little affected. As a general measure of the protection associated with the answer $Y=y$ we could take $V(X|Y=y)$, and as an overall measure of the protection provided by an RR-scheme we could take the average $EV(X|Y)$. Alternatively, we could use the relative measures $V(X|Y=y)/V(X)$ and $EV(X|Y)/V(X)$.

The quantity $EV(X|Y)$ is very natural from a statistical point of view. One could thus anticipate that it might be related to the loss of information $I(X|Y)$ in some simple way. In Examples 1 and 2, where θ is the parameter of a binomial distribution and the mean of a normal distribution, respectively, simple relations can be established. Further examples will be given in Section 5.

EXAMPLE 1 (continued) From Example 1, p. 8, we have

$$I(X|Y=y) = \frac{h_Y(y|1)h_Y(y|0)}{\theta(1-\theta)[\theta h_Y(y|1) + (1-\theta)h_Y(y|0)]^2}$$

and from p.7

$$P(X=1|Y=y;\theta) = 1 - P(X=0|Y=y;\theta) = \frac{\theta h_Y(y|1)}{\theta h_Y(y|1) + (1-\theta)h_Y(y|0)}.$$

Hence the conditional information given $Y=y$ can be written

$$\begin{aligned} I(X|Y=y) &= \frac{P(X=1|Y=y;\theta)P(X=0|Y=y;\theta)}{\theta^2(1-\theta)^2} = \\ &= V(X|Y=y)/\theta^2(1-\theta)^2, \end{aligned}$$

and, taking the average with respect to $f_Y(y;\theta)$,

$$I(X|Y) = EV(X|Y)/\theta^2(1-\theta)^2.$$

The loss of information thus is proportional to the protection as measured by $EV(X|Y)$. Compare also Remark 2, p.17.

EXAMPLE 2 (continued) From Example 2, p.8, we directly have

$$V(X|Y=y) = \sigma^2/(1+\sigma^2)$$

and

$$I(X|Y=y) = \sigma^2/(1+\sigma^2).$$

Thus, by taking the average with respect to $f_Y(y;\theta)$, we have

$$I(X|Y) = EV(X|Y)$$

and so the loss of information equals the protection as measured by $EV(X|Y)$.

Hitherto the special background of RR has not been considered. The aim of RR is to protect the interviewees in order to make them more co-operative, thereby reducing the answer bias. In some situations (e.g. surveys concerning income or political sympathy) every value of X might be embarrassing. Then $V(X|Y=y)$ is a relevant measure of the protection given $Y=y$. However, in most situations only some of the X -values are sensitive (e.g. surveys concerning illegal abortion). Let these values constitute the set A . Then the protection of an RR-scheme should be directed towards this set, and the conditional probability of A given $Y=y$, $P(A|Y=y;\theta)$, might be a relevant measure of the protection given $Y=y$.

When judging the protection provided by a whole RR-scheme every possible answer Y must be considered. Suppose the randomizing device has produced the answer $Y=y$. Before delivering this answer to the interviewer, the respon-

dent might consider the suspicions that this answer will throw upon him, and if he thinks that they are too stigmatizing, he might choose to lie or refuse to answer. Hence it is natural to demand a minimal protection from every possible answer and so we are led to measures of the types

$$\min_y V(X|Y=y)$$

and

$$\max_y P(A|Y=y;\theta).$$

These measures of the protection are more realistic than $EV(X|Y)$ to measure the answer-bias reducing effects of an RR-scheme. The following examples show how they can be used in planning situations.

EXAMPLE 1 (continued) Let the characteristic A correspond to "tax evader".

Then we have

$$P(A|Y=y;\theta) = P(X=1|Y=y;\theta) = \frac{\theta h_Y(y|1)}{\theta h_Y(y|1) + (1-\theta)h_Y(y|0)}.$$

In Anderson (1975a) the following condition is considered:

$$(4.1) \quad \max_y P(A|Y=y;\theta) \leq \text{constant}$$

for some fixed θ believed to be approximately correct. It is shown that, subject to this condition, plans with essentially two possible answers yield the most efficient estimates of θ .

EXAMPLE 2 (continued) The embarrassing variable $X \in N(\theta, 1)$. Allow an arbitrary family of response densities, $\{h_Y(y|x); -\infty < x < \infty\}$. We shall in this example determine a family of response densities that minimizes the loss of information $I(X|Y)$ subject to the condition

$$(4.2) \quad \min_y V(X|Y=y) \geq \sigma^2/(1+\sigma^2),$$

where σ^2 is a known constant. In Section 5, Remark 2, it is shown that $I(X|Y) = EV(X|Y)$. Hence by taking the average over Y in (4.2) we have

$$(4.3) \quad I(X|Y) \geq \sigma^2/(1+\sigma^2).$$

But the special family of response densities $\{N(x, \sigma); -\infty < x < \infty\}$ of Example 2, p.4, fulfils condition (4.2) (see Example 2, p.13). Moreover, equality holds in (4.3) for this particular family (see Example 2, p.8). Hence

the family $\{N(x, \sigma); -\infty < x < \infty\}$ is an optimal family of response distributions subject to the condition (4.2).

So far we have tried to summarize in a single number the protection provided by an RR-scheme. Generally the privacy aspects are too complex to allow such a compression. This is the case for instance in surveys concerning political sympathy. We could then introduce a multi-dimensional measure of the protection in the following way. Divide the sample space of X into k classes, $A_1, A_2, \dots, A_k$. The vector $P(A|Y=y;\theta) = (P(A_1|Y=y;\theta), \dots, P(A_k|Y=y;\theta))$ then measures the protection associated with the answer $Y=y$. In analogy with Condition (4.1) in Example 1, p.14, one could demand that the vector $P(A|Y=y;\theta)$ should assume values in some non-embarrassing set Q_θ for every possible answer y and for some fixed θ believed to be approximately correct. By choosing the classes $A_1, \dots, A_k$ and the set Q_θ appropriately, most demands of protection can be handled. This approach is worked out in Anderson (1975a). For example, it is shown that there exists an RR-scheme with at most k possible answers that maximizes the information $I(Y)$ subject to conditions of the above type.

In Section 5 we shall study two examples of special RR-models. The first example deals with exponential families of densities of X and the second example with the "given-answer model". In both cases the aim is to illustrate the conflict between efficiency and protection.

5. TWO EXAMPLES

5.1. Exponential families.

Most probability models that are used in practice belong to an exponential family. With the natural parameterization the density of X can then be written

$$f_X(x; \theta) = k(x) \beta(\theta) \exp\{\sum_1^r \theta_i t_i(x)\}$$

with respect to some measure $d\mu(x)$. Assume, first, that θ is one-dimensional. Then $f_X(x; \theta) = k(x) \beta(\theta) \exp\{\theta t(x)\}$ and Fisher's information is given by

$$I(X) = -E \left[\frac{\partial^2 \log f_X(X; \theta)}{\partial \theta^2} \right] = - \frac{d^2 \log \beta(\theta)}{d\theta^2} = V(t(X));$$

for the last equality, see Ferguson (1967), p.131. Suppose as before that $\{h_Y(y|x); x \in \Omega_X\}$ is the family of response densities. The observable answer Y from a randomly taken individual has the density

$$f_Y(y; \theta) = \beta(\theta) \int h_Y(y|x) k(x) \exp\{\theta t(x)\} d\mu(x) = \beta(\theta) A(y; \theta).$$

Fisher's information based on Y is

$$\begin{aligned} I(Y) &= -E \left[\frac{d^2 \log \beta(\theta)}{d\theta^2} + \frac{\partial^2 \log A(Y; \theta)}{\partial \theta^2} \right] = \\ &= I(X) - E \left[\frac{\partial^2 \log A(Y; \theta)}{\partial \theta^2} \right]. \end{aligned}$$

Hence, from Theorem 1 the conditional information from X given Y

$$I(X|Y) = -E \left[\frac{\partial^2 \log A(Y; \theta)}{\partial \theta^2} \right].$$

This expression can be simplified. The relations

$$\frac{\partial^2 \log A(y; \theta)}{\partial \theta^2} = \frac{\partial^2 A(y; \theta)}{\partial \theta^2} / A(y; \theta) - \left(\frac{\partial A(y; \theta)}{\partial \theta} \right)^2 / (A(y; \theta))^2,$$

$$\frac{\partial A(y; \theta)}{\partial \theta} = \int h_Y(y|x) k(x) t(x) \exp\{\theta t(x)\} d\mu(x),$$

$$\frac{\partial^2 A(y; \theta)}{\partial \theta^2} = \int h_Y(y|x) k(x) (t(x))^2 \exp\{\theta t(x)\} d\mu(x)$$

yield

$$\frac{\partial^2 \log A(y; \theta)}{\partial \theta^2} = E(t^2(X) | Y=y) - (E(t(X) | Y=y))^2 = V(t(X) | Y=y).$$

Hence we have

THEOREM 3a Let the density $f_X(x;\theta)$ of the embarrassing variable X belong to a one-dimensional exponential family, i.e. $f_X(x;\theta) = k(x)\beta(\theta)\exp\{\theta t(x)\}$ with respect to some measure $d\mu(x)$. Then the decrease in Fisher's information due to randomized response can be written

$$I(X|Y) = EV(t(X)|Y).$$

REMARK 1 Since $I(X) = V(t(X))$ we also have

$$I(Y) = V(t(X)) - EV(t(X)|Y) = VE(t(X)|Y).$$

REMARK 2 In the very special situation when $t(x) = x$ we have

$I(X|Y) = EV(X|Y)$, which is sometimes a relevant measure of the protection (see Section 4).

For an r -dimensional parameter θ ($r > 1$) we have analogously to the one-dimensional case:

THEOREM 3b Let the density $f_X(x;\theta)$ of the embarrassing variable X belong to an r -dimensional exponential family, i.e. $f_X(x;\theta) = k(x)\beta(\theta)\exp\{\sum_{i=1}^r \theta_i t_i(x)\}$ with respect to some measure $d\mu(x)$. Then the decrease in Fisher's information matrix $I(X|Y)$ has the elements

$$I_{ij}(X|Y) = ECov(t_i(X), t_j(X)|Y), \quad i, j = 1, 2, \dots, r.$$

5.2. The model with given answers

In this example we shall consider the model with given answers proposed by Eriksson (1973) (compare the scheme of Greenberg et al. (1971)). The respondents then answer with their true values of the sensitive characteristic X with probability p and according to a known probability density $g(y)$ with probability $1-p$. The answer Y from a randomly chosen individual has the density

$$f_Y(y) = pf_X(y;\theta) + (1-p)g(y).$$

Assume that f_X and g are densities of continuous random variables,

that $\{f_X(x;\theta); \theta \in \Theta\}$ is a regular family, and that θ is one-dimensional. If a respondent answers $Y=y$, then this answer could be either his true value or it could emanate from the known density $g(y)$. Since X and Y are continuous random variables the conditional probability that y is the true value is

$$(5.1) \quad P(X=y|Y=y) = \frac{pf_X(y;\theta)}{pf_X(y;\theta) + (1-p)g(y)}.$$

Fisher's information from an observation Y is

$$I(Y) = E \left[\frac{\partial \log f_Y(Y;\theta)}{\partial \theta} \right]^2 = \int p^2 \left[\frac{\partial}{\partial \theta} f_X(y;\theta) \right]^2 / (pf_X(y;\theta) + (1-p)g(y)) dy$$

or after substituting (5.1)

$$I(Y) = p \int P(X=y|Y=y) \left[\frac{\partial}{\partial \theta} f_X(y;\theta) \right]^2 / f_X(y;\theta) dy.$$

This expression differs from the information from a direct observation X

$$I(X) = \int \left[\frac{\partial}{\partial \theta} f_X(y;\theta) \right]^2 / f_X(y;\theta) dy$$

in two respects. The factor p can be explained by the fact that only a proportion p of the respondents answer with their true values. The factor $P(X=y|Y=y)$ in the averaging integral measures the degree of protection for different answers y .

Let us consider some planning aspects. If every x -value might be stigmatizing it is reasonable to require that

$$(5.2) \quad P(X=y|Y=y) \leq k$$

for every possible answer y with the constant k chosen under the circumstances. From (5.1) we have that this is equivalent to

$$k(1-p)g(y) \geq p(1-k)f_X(y;\theta).$$

After integration we see that $p \leq k$. If $g(y)$ is chosen equal to $f_X(y;\theta)$ when designing the RR-scheme (in practice one must guess) and p is taken equal to k , then $P(X=y|Y=y) = k$ and we see that this plan makes the information $I(Y) = k^2 I(X)$ maximum, subject to Condition (5.2).

Often the density of the sensitive characteristic $f_X(x)$ is not parametric. The model with given answers still allows an unbiased estimate of the mean $m_f = E(X)$:

$$m_f^* = (\bar{y} - (1-p)m_g)/p$$

with variance

$$V(m_f^*) = V(Y)/np^2 = \\ = [p\sigma_f^2 + (1-p)\sigma_g^2 + p(1-p)(m_f - m_g)^2]/np^2,$$

where $\sigma_f^2 = V(X)$, $m_g = \int yg(y)dy$ and $\sigma_g^2 = \int (y - m_g)^2 g(y)dy$ and n is the number of observations. Taking $g(y) = f_X(y)$ and $p = k$ we get

$$P(X = y | Y = y) = k$$

and

$$V(m_f^*) = V(X)/nk^2$$

to be compared with the above relation $I(Y) = k^2 I(X)$.

REFERENCES

- Abul-Ela, A.-L.A., Greenberg, B.G., Horvitz, D.G. (1967). A multi-proportions randomized response model. *J. Amer. Statist. Assoc.* 62, 990-1008.
- Anderson, H. (1975a). Efficiency versus protection in randomized response designs for estimating proportions. Technical report, Department of Mathematical Statistics, University of Lund.
- Dalenius, T., Vitale, R.A. (1974). A new randomized response design for estimating the mean of a distribution. Report no. 78 of the research project Errors in Surveys, Institute of Statistics, University of Stockholm.
- Eriksson, S. (1973). Randomized Interviews for Sensitive Questions. Ph.D. Thesis, Institute of Statistics, University of Göteborg.
- Ferguson, T.S. (1967). *Mathematical Statistics. A Decision Theoretic Approach.* Academic Press, New York.
- Fraser, D.A.S. (1965). On information in statistics. *Ann. Math. Statist.* 36, 890-896.
- Greenberg, B.G., Abul-Ela, A.-L.A., Simmons, W.R., Horvitz, D.G. (1969). The unrelated question randomized response model: Theoretical framework. *J. Amer. Statist. Assoc.* 64, 520-539.
- Greenberg, B.G., Kuebler Jr, R.R., Abernathy, J.R., Horvitz, D.G. (1971). Application of the randomized response technique in obtaining quantitative data. *J. Amer. Statist. Assoc.* 66, 243-250.
- Kullback, S. (1959). *Information Theory and Statistics.* Wiley, New York.
- Papaioannou, P.C., Kempthorne, O. (1971). On statistical information theory and related measures of information. Technical report no. ARL 71-0059, Aerospace Research Laboratories, Wright-Patterson AFB, Ohio 45433.
- Poole, W.K. (1974). Estimation of the distribution function of a continuous type random variable through randomized response. *J. Amer. Statist. Assoc.* 69, 1002-1005.
- Warner, S.L. (1965). Randomized response: A technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.* 60, 63-69.
- Warner, S.L. (1971). The linear randomized response model. *J. Amer. Statist. Assoc.* 66, 884-888.

